

MIND THE GAP

Miksi fysikalismi ei selitä tietoisuutta

Pii Telakivi

Helsingin yliopisto
Humanistinen tiedekunta
Teoreettinen filosofia
Pro gradu -tutkielma
Huhtikuu 2012

Sisältö

1. Johdanto	1
1.1. Tietoisuuden käsitteestä	4
1.2. Fysikalismien käsitteestä	8
2. Selityksellinen argumentti	11
2.1. What is it like -argumentti	13
2.2. Selityksellinen kuilu	15
2.3. Selityksellisen argumentin ontologinen muoto	20
2.4. Yhteenveto selityksellisestä argumentista	23
3. Modaaliargumentti	24
3.1. Kripken psykofyysisten identiteettien kritiikki	26
3.2. Puuttuvan kvalian argumentti	29
3.3. Käänteisen kvalian argumentti	31
3.4. Seuraako kuviteltavuudesta metafyyminen mahdollisuus?	33
3.5. Moninaisen toteutuvuuden argumentti	39
3.6. Modaaliargumentin tulokset	44
4. Tietoargumentti	46
4.1. Jacksonin tietoargumentin fysikalismi	48
4.1.1. Kysymys 1: Oliko Marylla todella kaikki fysikaalinen tieto ennen vapautumistaan?	52
4.1.2. Kysymys 2: Oppiiko Mary jotakin uutta?	53
4.1.3. Kysymys 3: Onko oppiminen faktuaalista vai ei-faktuaalista?	56
4.1.4. Kysymys 4: (Jos ei-faktuaalista, niin) saako hän vain uuden taidon?	57
4.1.5. Kysymys 5: (Jos ei-faktuaalista, niin) saako hän vain tuttuustietoa?	60
4.1.6. Kysymys 6: (Jos faktuaalista, niin) saako hän tietoonsa uusia faktoja, vai vain vanhoja faktoja uudella tavalla ymmärrettynä?	62
4.1.7. Kysymys 7: (Jos faktuaalista, niin) kumpi erottelutapa on oikea?	64
4.1.8. Kysymys 8: Kumoaako Maryn sekundaarisella tavalla oppima uusi fakta fysikalismien?	65
4.2. Johtopäätökset kahdeksasta kysymyksestä	66
5. Fysikalistin puolustuspuheenvuoro	69
6. Lopuksi	75
Lähteet:	77

How is it that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of Djin when Aladdin rubbed his lamp.¹

1. Johdanto

H. G. Wellsin novellissa *The Country of the Blind* (1904) vuorikiipeilijä putoaa satojen metrien pudotuksen jyrkänteeltä Andien keskellä. Muu seurue luulee hänen menehtyneen, mutta pehmeään lumeen putoaminen säästää hänen henkensä. Hän erottaa kaukana alapuolellaan mäntyjä kasvavan laakson, jonne hän pääsee vaikeakulkuisen rotkon läpi. Laakso ei olekaan asumaton, vaan siellä asuu sokeiden yhdyskunta, joka on sukupolvien ajan ollut täysin eristyksissä muusta maailmasta. He elävät laamoja kasvattaen laakson suotuisassa ilmastossa. Perimätiedon mukaan satoja vuosia sitten laakson asukkaita alkoi vaivata erikoinen sairaus, jossa he menettivät näkönsä vähitellen: vanhat hapuilivat ja siristelivät, nuoremmat näkivät hämärästi, ja vastasyntyneet olivat syntymästään saakka sokeita. Sokeutuminen tapahtui kuitenkin niin hitaasti, että ihmiset ehtivät oppia tarvittavat käytännön taidot ilman näköäkin. Vuodet vierivät, ja yhdyskunta sopeutui niin hyvin sokeaan elämään, että heidän silmänsä surkastuivat ja lopulta koko näkemisen käsitekin unohtui.

Vuorikiipeilijän huomattua päätyneensä Sokeiden Valtakuntaan, hän kertoo ensimmäisille tapaamilleen ihmisille, että hän tulee suuresta maailmasta vuorten takaa Bogotan läheltä – ja että hän näkee. Sokeat eivät ymmärrä selostuksesta kuitenkaan muuta kuin sanan ”Bogota”. Tämä hullu mies siis ilmeisesti yrittää selittää olevansa nimeltään Bogota. Mitkään todistusyritykset ylimääräisestä aistista eivät auta: sokeiden kielestä ovat jääneet pois merkityksettömät näkemiseen ja näkökykyyn liittyvät sanat, ja he pitävät Bogotaa kehittymättömänä hourailevana alkuasukkaana.

¹ T.H. Huxley (1866): *Lessons in Elementary Psychology*, s. 210. Lainattu teoksesta Block 2002, 207.

Onko Bogotan mahdollista saada sokeat ymmärtämään, mitä on näkeminen? Riittävätkö mitkään mahdolliset *selitystavat*, vai vaatiiko näkemisen ymmärtäminen välttämättä subjektiivisen kokemuksen – näkemisen?

Osoitan tässä pro gradu -työssäni kolmen argumentin avulla mielenfilosofisen fysikalismin ongelmallisuuden. Argumentit kritisoivat fysikalismia eri tavoin, mikä mahdollistaa monipuolisen katsauksen fysikalismin ongelmakohtiin. **Selityksellinen argumentti** kysyy voiko tietoisuuden *selittää* fysikaalisin termein – tässä argumentissa kuilu aivojen ja tietoisuuden välissä on siis episteeminen. **Modaaliargumentti** kysyy pikemminkin *onko* tietoisuus fysikaalinen, eli sen esittämä kuilu on ontologinen. **Tietoargumentti** lähtee episteemisestä kuilusta, joka johtaa ontologiseen kuiluun. Epistemologinen ja ontologinen kysymys ovat kuitenkin läheisessä yhteydessä toisiinsa, ja sekoittuvat jokaisessa argumentissa. David Chalmers esimerkiksi johtaa selityksellisestä argumentista ontologisen johtopäätöksen. Modaaliargumenttia voi käyttää episteemisenä argumenttina, jos pitää systeemin fysikaalista kopiota ilman tietoisuutta *loogisena mahdollisuutena*. Toisaalta sitä voi pitää myös ontologisena argumenttina, jos pitää fysikaalista kopiota ilman tietoisuutta *metafyysisenä mahdollisuutena*. Frank Jackson esittää tietoargumentin ontologisena argumenttina, mutta tietoargumentin kriitikot saattavat hyväksyä tietoargumentista vain episteemisen version. Puhun siis *episteemisestä argumentista*, kun väitetään, että fysikalismi ei selitä tietoisuutta, ja *ontologisesta argumentista*, kun väitetään, että tietoisuus ei ole (kokonaan) fysikaalinen.

Kaikissa kolmessa argumentissa osoitetaan ensin, että tietoisuuden ja fysikaalisten faktojen välissä on episteeminen kuilu. Ensimmäinen argumentti keskittyy selittämiseen, toinen kuvittelemiseen ja kolmas tietämiseen. Olen jakanut kaksi ensimmäistä argumenttia alalajeihin, jotka ovat näkökulmiltaan tai lähtökohdiltaan jollain tavalla toisistaan poikkeavia. Kolmannen argumentin taas olen jakanut osaluokkiin kritiikkiensä perusteella. Työni perimmäisenä tarkoituksena ei ole tutkia näiden argumenttien itsessään pätevyyttä, vaan käyttää niitä pikemminkin apuvälineinä osoittamaan, että reduktiivinen fysikalismi on väärässä. Mutta jotta ne kumoaisivat fysikalismin, täytyy niiden tietenkin olla kestäviä argumentteja. Sen takia pyrin tutkimaan laajasti näiden argumenttien itsessään pätevyyttä. Lopuksi

tarkastelen vielä tapoja, joilla fysikalistit voivat yrittää puolustautua näitä argumentteja vastaan.

Väitän, että argumentit osoittavat vakuuttavasti, että fysikaalisen aivotoiminnan ja tietoisuuden välissä on aukko, jota (ainakaan reduktiivinen) fysikalismi ei selitä.² Kaikki argumentit yhdessä osoittavat tämän, vaikka vain yksi (selityksellinen argumentti, kpl 2) on eksplisiittisesti muotoiltu tämän osoittamiseen. Lopputulokseni – fysikalistisen selityksen kyvyttömyys tietoisuuden suhteen – osoittaa tietynlaisen mielenfilosofian (so. eliminatiivisen materialismin ja ylipäätään reduktiivisen fysikalismin) olevan lähtökohdiltaan väärässä. Fysikalismi ei voi selittää tietoisuutta samalla tavalla kuin se voi selittää kinestetiikan tai osmoosin. Tämän takia teorit, jotka perustuvat tietoisuuden selittämiseen samoilla keinoilla kuin muiden maailman ilmiöiden selittämiseen, on tuomittu epäonnistumaan. Useat filosofit ovat päätyneet kanssani samaan ratkaisuun – mutta päätyneet siitä eteenpäin erilaisiin johtopäätöksiin. Osa on sitä mieltä, että fysikalismin kyvyttömyydestä selittää tietoisuus seuraa, että fysikalismi on väärässä; tietoisuus (tai sen osa) koostuu jostain ei-fysikaalisesta. Osa taas on sitä mieltä, että selityskyvyttömyydestä huolimatta tietoisuus on pohjimmiltaan (ontologisesti) fysikaalinen.

Kysymys onkin siis seuraavanlainen: mitä siitä seuraa, että fysikalismi ei selitä tietoisuutta? Mahdollisuuksia on ainakin kolme. 1. Fysikalismi on väärässä, ja on hyväksyttävä jonkinlainen dualismi. 2. Fysikalismi ei ole väärässä, sillä kaikki on kyllä pohjimmiltaan fysikaalista – ongelma on ainoastaan selityksellinen siinä mielessä että a) emme vain ole vielä kyenneet selittämään sitä b) emme koskaan kykene selittämään sitä käsityskykymme rajoittuneen luonteen takia c) mikään muukaan kognitiivinen järjestelmä ei koskaan periaatteessakaan voi selittää sitä. 3. Koko fysikalismin käsite on jotenkin väärin muodostettu, ja sen takia se ei kykene selittämään tietoisuutta (väärään kysymykseen ei voi saada oikeaa vastausta). Toinen ja kolmas vaihtoehto sopivat dualismia paremmin muuhun käsitykseen todellisuuden rakenteesta, joten on kenties jokseenkin perusteltua asettua toisen niistä kannalle.

² Se, onko tämä ongelmallista fysikalismille, on asia erikseen. Esim. Van Gulick (2003) on esittänyt, että selityksellinen kuilu onkin oikeastaan *selittävä kuilu*, joka selittää tietoisuuden luonnetta sekä omaa olemassaoloaan. Kuilu ei heikennä fysikalismin asemaa, vaan päinvastoin se on täysin odotettavissa, ja vain selventää subjektiivisen ja objektiivisen erilaista luonnetta.

1.1. Tietoisuuden käsitteestä

Without consciousness the mind-body problem would be much less interesting. With consciousness it seems hopeless. (Nagel 1974, 436.)

Tietoisuus on epäilemättä mielenfilosofian suurin ratkaisematon haaste, joka tuntuu pakenevan selityksiä kuin horisontti tavoittelijaansa. Erityisen vaikean siitä tekee siihen kuuluva fenomenaalinen puoli – se että kokemukset ovat juuri sellaisia kuin ovat eli *jonkinlaisia*. Eläin, ihminen tai muu kognitiivinen systeemi voi olla tietoinen monessa mielessä: olemalla valveilla, tietoinen itsestään ja ympäristöstään, tietoinen joistain faktoista, kykenevä introspektioon sekä kommunikoimaan siitä muille, olemalla tietoisten tilojensa subjekti jne. Esimerkiksi Ned Block (2002) on tehnyt mielestäni selventävän erottelun jakamalla tietoisuuden käsitteen neljään alalajiin. Tietoisuudesta puhuttaessa on syytä kiinnittää erityistä tarkkuutta siihen, millä tavalla käsitettä käyttää, sillä lopputulokset saattavat vaihdella suuresti riippuen siitä millaisesta tietoisuuden käsitteestä on kyse. Blockin (2002, 208–214) jaottelu tietoisuuden käsittämisen tavoista:

1. Fenomenaalinen tietoisuus (F-tietoisuus) on kokemuksellisuutta, jossa olennaista on se, että on *jonkinlaista* (something *it is like*) (vrt. Nagel 1974) olla tietyssä tilassa. Kun kuulemme, näemme, haistamme, maistamme tai tunnemme kipua, on kyse fenomenaalista (eli kvalitatiivista) tietoisuuden tiloista. Sen ominaisuudet ovat kokemuksellisia: tuntemukset, aistimukset ja havainnot sekä Blockin (s. 206) mukaan myös ajatukset, halut ja tunteet. Block ei puhu *kvalioista*, mutta kyseessä on hyvin pitkälti sama asia. Kvaliat (*qualia*) ovat fenomenaalisen tietoisuuden ominaisuuksia, toisin sanoen aistimellisia tiloja, jotka määräävät *minkälaista* on olla kyseisessä tilassa. Kvalioita kuvataan usein *raakatuntemuksiksi* (*raw feels*). Joskus puhutaan vain *fenomenaalista ominaisuuksista* tai *kvalitatiivisista piirteistä*. Esimerkiksi punaisuuden punaisuus, sitruunan kirpeys sekä kivun kivuliaisuus ovat kvalioita.

2. Pääsy tietoisuudessa (P-tietoisuus, *access consciousness*) subjekti on tietoinen, kun sillä on esteetön pääsy sisäisiin tiloihinsa. Pääsy tietoisuus on sekä representationaalista että funktionaalista, mutta ei fenomenaalista kuten F-tietoisuus.

Kyse ei siis ole siitä, onko tietoisuus *jonkinlaista*, vaan siitä, että tietoisella organismilla on hallussaan tarvittava informaatio, joka antaa valmiuden puheeseen, päättelyyn ja toimintaan. Tällaisia P-tyypin tietoisuuden tiloja ovat *propositionaaliset asenteet*, toisin sanoen ajatukset, uskomukset ja halut – tilat joiden representationaalisen sisällön voi ilmaista ”että”-lauseilla. (”Ajatus, että ruoho on vihreää.”)

3. Itsetietoisuuteen (I-tietoisuus, *self-consciousness*) kuuluu kyky käyttää *itsen* käsitettä itsestään. Se on siis tietoisuutta siitä, että se on juuri kyseinen *itse*, joka on tietoinen itsestään.³ F-tietoisuus on myös usein *itse*en liittyvää, sillä fenomenaaliset tilat (kuten kipu, lämmöntunne, keltaisen keltaisuus jne.) ovat subjektin, itsen, eli *minun* tiloja. F-tietoisuus ja I-tietoisuus eivät kuitenkaan ole sama asia, sillä *itse* pysyy samana, vaikka fenomenaaliset sisällöt muuttuvat. Punaisuuden ja vihreiden kokemukset ovat I-tietoisuuden kannalta sama asia, mutta F-tietoisuuden kannalta eri asia.

4. Tarkkailevassa tietoisuudessa (T-tietoisuus, *monitoring-consciousness*) tietoisuus voidaan ensinnäkin määritellä jonkinlaiseksi sisäiseksi havainnoinniksi henkilön omista tiloista tai itsestä. (Tässä mielessä se voi myös samalla olla F-tietoisuutta tai I-tietoisuutta.) Toisaalta tarkkailevan tietoisuuden voi ymmärtää sisäiseksi tutkiskeluksi informaation prosessoinnin mielessä (*internal scanning*). Eliminativistit voivat redusoida F-tietoisuudenkin tällaiseksi sisäiseksi prosessoinniksi. Vaikka T-tietoisuus voi joskus olla F-tietoisuutta, tämä tietoisuuden muoto on syytä erottaa F-tietoisuudesta: esim. tietokoneella voi olla T-tietoisuutta, mutta ei (oletettavasti) F-tietoisuutta.

Blockin mukaan (2002, 208) tietoisuuden kaksi ensimmäistä lajia, F- ja P-tietoisuus, sekoitetaan usein keskenään, ja siitä seuraa vakavia sekaannuksia. Hänen mukaansa on kuviteltavissa tilanteita, joissa organismilla on P-tyypin tietoisuutta, mutta ei F-

³ On vaikeaa jaotella tarkkaan, millä eläinlajeilla tätä tietoisuuden tapaa esiintyy, mutta on syytä otaksua, että ihmisten lisäksi sitä esiintyy ainakin jonkinlaisessa muodossa eräillä muilla kädellisillä. Block käyttää kriteerinä sitä, tunnistaako eläin itsensä peilistä. Simpanssi tunnistaa itsensä, koira ei. Block kuitenkin huomauttaa, että vaikka koirat eivät olisikaan itsetietoisia, niillä voi silti olla fenomenaalisia tietoisia tiloja, toisin sanoen niistä tuntuu *jonkinlaiselta* katsella kuvaansa peilistä. (Block 2002, 214.) Itse pitäisin peilistä tunnistamista liian ihmiskeskeisenä kriteerinä, joka esimerkiksi korostaa liikaa näköaistin merkitystä.

tyypin tietoisuutta. *Zombi* (fysikaalisesti ihmisenkaltainen, mutta ilman fenomenaalista tietoisuutta oleva olento) on täydellinen esimerkki tällaisesta tietoisuudesta, mutta koska zombien käsitteellisestä (ja varsinkin metafyyysisestä) mahdollisuudesta vallitsee kaikkea muuta kuin yksimielisyys, ja zombi-argumentti johtaa itsessään moniin lisäkysymyksiin, on tässä tapauksessa suotavampaa käyttää muuta esimerkkiä.⁴ Blockin (2002, 211) maanläheisempi esimerkki on sokeanäkö, joka toimii kuin osittainen zombius: siitä kärsivät henkilöt osaavat erottaa (”arvata”) esimerkiksi kuvioita toisistaan, vaikka heillä ei olekaan tietoista kokemusta niiden näkemisestä. Tällainen arvaaminen on kuitenkin melko rajattua: potilaille pitää antaa selkeät vaihtoehdot (esim. ”X/O”, ”vaakasuora/pystysuora”), jotta he onnistuvat ”arvauksessaan”. Block kehottaa kuvittelemaan eräänlaisen *supersokeanäön*: tilanteen, jossa sokeanäköpotilas on harjoittanut itseään asettamalla itselleen vaihtoehtoja, mitä sokealla alueella saattaisi olla. Tällä tavalla hän saa tietoonsa näkökentässään olevia asioita, vaikka ei näekään niitä.

The superblindsighter himself contrasts what is it like to know visually about an 'X' in his blind field and an 'X' in his sighted field. There is something it is like to experience the latter, but not the former he says. It is the difference between *just knowing* and knowing via a visual experience. (Block 2002, 211.)

Vaikka tällaista superkykyä ei (oletettavasti) maailmassa esiinny, Blockin mukaan se osoittaa kuitenkin, että on mahdollista olla P-tietoisuutta ilman F-tietoisuutta. Itse sanoisin, ettei tarvitse postuloida mitään superkykyä, vaan jo olemassa oleva sokeanäkö todistaa asian aivan samalla tavoin.

Myös David Chalmers (1996) tarkentaa tietoisuuden käsitettä vastaavalla tavalla, hän tosin puhuu vain kahdesta tietoisuuden lajista: **1. fenomenalisesta tietoisuudesta** (joka vastaa Blockin fenomenaalista tietoisuutta) ja **2. psykologisesta tietoisuudesta** (johon kuuluvat muut Blockin tietoisuuden lajit). Erottelu on siis pääpiirteissään sama kuin Blockilla. Tarkalleen ottaen Chalmersin (1996, 26–27) määrittelemänä psykologiseen tietoisuuteen kuuluu: *valveillaolo* (erotuksena nukkumisesta),

⁴ Käsittelen zombi-argumenttia kappaleessa 3.2.

introspektio (pääsy omiin mentaalisiin tiloihin), *raportoitavuus*⁵ (vaatii introspektiota sekä sen lisäksi kykyä käyttää kieltä), *itsetietoisuus*, *tarkkaavaisuus* (joku on tietoinen jostain kiinnittäessään siihen huomiota), *tahdonalainen kontrolli* (teko on tietoinen kun se on tehty vapaaehtoisesti) ja *tieto* (arkipuheessa sanotaan usein, että on *tietoinen* jostain faktasta, silloin kun yksinkertaisesti vain *tietää* tuon faktan). Chalmers (1995; 2002a) kutsuu psykologisen tietoisuuden kysymyksiä myös *tietoisuuden helpoiksi ongelmiksi*,⁶ sillä ne ovat Chalmersin (samoin kuin Blockin) mukaan selitettävissä funktionaalisin termein.⁷ Sen sijaan fenomenaalista tietoisuutta funktionalismi ei kykene selittämään. Fenomenaalista tietoisuutta (eli fenomenaalista kokemusta) Chalmers kutsuukin *tietoisuuden vaikeaksi ongelmaksi*.

What makes the hard problem hard? Here, the task is not to explain behavioral and cognitive functions: even once one has an explanation of all the relevant functions in the vicinity of consciousness – discrimination, integration, access, report, control – there may still remain a further question: why is the performance of these functions accompanied by experience? Because of this, the hard problem seems to be a different sort of problem, requiring a different sort of solution. (Chalmers 2002a, 248.)

Psykologiseen pääsy tietoisuuteen kuuluvista asioista ollaan tietoisia siinä mielessä, että niistä ollaan perillä, valveutuneita (*aware of*). Chalmersin (1996, 28) mukaan tämänlaista tietoisista kokemusta kuvaa parhaiten juuri termi ”valveutuneisuus” (*awareness*). ”Awareness” ja ”consciousness” -termeistä ei ole suomen kielessä vastaavaa sanaparia, joten kutsun niitä tässä a-tietoisuudeksi ja c-tietoisuudeksi. Hänen mukaansa on tärkeää erottaa a-tietoisuus c-tietoisuudesta: c-tietoisuuteen kuuluu aina myös a-tietoisuutta, esimerkiksi fenomenaaliseen kokemukseeni pöydällä olevasta punaisesta kirjasta kuuluu välttämättä funktionaalinen a-tietoisuus sekä kirjasta että punaisesta väristä. A-tietoisuutta sen sijaan voi olla ilman c-

⁵ Raportoitavuuden vaatimus tulee esiin Sokeiden Valtakunnan laamoissa: laamoilla on aina ollut näkökyky, muttei mitään keinoa raportoida siitä sokeille. (Toisaalta sokeillakaan ei olisi koskaan ollut mitään keinoa raportoida laamoille sokeudestaan, vaikka he olisivatkin tienneet siitä.)

⁶ ”Helppous” on Chalmersin termissä toki suhteellisessa käytössä: kognitiotieteellä ja neuropsykologialla saattaa vierähtää vuosisata jos toinenkin kysymysten ratkaisemisessa, mutta ei ole mitään periaatteellista syytä, miksi ratkaisut eivät onnistuisi.

⁷ Kutsun selvyuden vuoksi tästä lähtien Chalmersin tietoisuuden helppoja ongelmia (eli psykologista tietoisuutta) sekä Blockin pääsy-, itse-, ja tarkkailevaa tietoisuutta yhteisnimityksellä *psykologinen pääsy tietoisuus*. (Siitäkin huolimatta, että I-tietoisuus ja T-tietoisuus voivat olla tiedostamattomia, toisin sanoen niihin ei ole välttämättä suoraa ”pääsyä”.)

tietoisuutta. Henkilö voi olla a-tietoinen jostain faktasta ilman fenomenaalista kokemusta, samalla tavoin kuin Blockin sokeanäkö-esimerkissä. Kaikki filosofit eivät kuitenkaan ole samaa mieltä fenomenaalisen tietoisuuden erityisasemasta, eivätkä varsinkaan siitä, voiko psykologista pääsytietoisuutta esiintyä ilman fenomenaalista tietoisuutta.

Tässä työssä käsitelty tietoisuus on juuri fenomenaalista tietoisuutta eli c-tietoisuutta (ellei toisin mainita). Käytän siitä ainakin nimityksiä ”*tietoisuus*”, ”*tietoinen kokemus*”, ”*subjektiivinen kokemus*”, ”*fenomenaalinen kokemus*” tai pelkästään ”*fenomenaalinen*”. Joskus saatan puhua vain ”*mentaalisesta*”, ja jotkut tässä työssä esittelemäni filosofit käyttävät niinkin epätarkkaa termiä kuin ”*psykologinen*”. Termit viittaavat kuitenkin yhteen ja samaan *fenomenaaliseen tietoisuuteen*. Monet käsittelemäni filosofit tuntuvat sekoittavan eri tietoisuuden tavat keskenään ja tulevat siitä johtuen antaneeksi vastauksia, jotka eivät vastaa alkuperäiseen kysymykseen.⁸

1.2. Fysikalismin käsitteestä

Kuten tietoisuus, myös fysikalismin (*physicalism*) käsite vaatii analysointia, sillä termillä voidaan ymmärtää hieman eri asioita näkökulmasta riippuen. Fysikalismi on monismin muoto, jonka mukaan *kaikki* on fysikaalista. Jos tietää kaiken fysikaalisen, tietää kaiken. ”*Fysikalismi*”-termiä käytetään usein synonyymisesti ”*materialismin*” kanssa, vaikka termeillä on omat taustansa ja erityispiirteensä. Fysikaaliset asiat ovat epäilemättä materiaalisia asioita.⁹ Fysikalismi kuitenkin implikoi enemmän kuin vain kohteensa materiaalisen luonteen: se asettaa *fysiikan* etusijalle todellisuuden selittämisessä. Kun halutaan tietää mitä on olemassa, tulee katsoa mitä fysiikka sanoo, että on olemassa. Siinä missä materialismi on sitoutunut a priori materian olemassaoloon, on fysikalismi sitoutunut materian olemassaoloon vain, jos fysiikka on sitoutunut materian olemassaoloon. Fysiikka-tiede koskee toki materiaalisia asioita, joten ”*materialismia*” ja ”*fysikalismia*” voi käyttää synonyymeina

⁸ Tämä tulee esiin varsinkin tietoargumentin kritiikeissä: useissa tapauksissa (esim. *taitotietohypoteesi*) kriitikko antaa mielestään tyhjentävän tietoisuuden selityksen, mutta tuleekin puhuneeksi aivan eri asiasta kuin tietoargumentissa oli kyse sekoittaessaan fenomenaalisen ja psykologisen pääsytietoisuuden.

⁹ Vaikka periaatteessa ne voisivat olla mitä tahansa asioita – tulevaisuuden fysiikkaan voisi kuulua vaikka immateriaalisia asioita, koska oppiin ei ole sisäänrakennettuna juuri materian vaatimusta, vaan vain fysiikka-tieteen vaatimus.

mielenfilosofisessa kontekstissa puhuttaessa tietoisuuden ja neurologian suhteesta. Tässä työssä käsitelty fysikalismi on materialismia, joten käytän tästä lähtien näitä kahta termiä synonyymisesti, ellei toisin mainita.

Nykyfysikalismi perustuu usein supervenienssiteoriaan, eli ajatukseen, että psykologiset/mentaaliset ominaisuudet määräytyvät fysikaalisten ominaisuuksien perusteella. Suhde toimii vain yhteen suuntaan: supervenienssiteorian mukaan asioiden, jotka ovat mentaalisesti samanlaisia, ei tarvitse välttämättä olla fysikaalisesti samanlaisia. Näkemys siis mahdollistaa tervejärkisen kannan, että fysikaalisesti kanssamme erilaiset oliot (vaikkapa avaruusoliot ja mustekalat) voivat kokea samoja mentaalisia tiloja, esim. kipua. Reduktiivinen fysikalismi taas väittää, että mentaaliset ominaisuudet määräytyvät fysikaalisten ominaisuuksien perusteella. Mieli redusoituu siis ruumiiseen: mentaaliset ominaisuudet ovat täysin riippuvaisia fysikaalisista ominaisuuksista. Mustekalat eivät siis voi tuntea samanlaista kipua kuin ihmiset, sillä niiden aivot ovat erilaiset. Fysikalistisista teorioista funktionalismiin kuuluu supervenienssi-ajattelu, ja identiteettiteoriaan reduktiivinen kanta.

Fysikalistille kaikki mitä on olemassa, *on* viimekädessä fysikaalista (tai muuten hän ei ole fysikalisti). Kantaa voisi kutsua *ontologiseksi fysikalismiksi*. Fysikalistin mukaan kaikkea ei välttämättä voi *selittää* fysikaalisin termein (mutta hän voi silti olla fysikalisti). Fysikalismia vastaan suunnattu kritiikki (kuten tässä työssä esittelemäni argumentit) on osoittanut fysikaalisen selityksen jättävän episteemisen kuilun, johon fysikalistin on reagoitava jotenkin. Yksi mahdollisuus on myöntää episteeminen kuilu, mutta kieltää, että siitä seuraisi ontologinen kuilu. Tällaisia kantoja ovat *non-reduktiivinen fysikalismi* ja *ominaisuusdualismi*. Niiden mukaan mentaaliset ominaisuudet eivät siis redusoidu fysikaalisiin ominaisuuksiin. Non-reduktiivinen fysikalismi hyväksyy supervenienssiteorian, mutta ominaisuusdualismi sen sijaan ei vaadi, että mentaaliset ominaisuudet olisivat välttämättä riippuvaisia fysikaalisista ominaisuuksista (Kim 2006, 290). Fysikalisti voi myös kieltää episteemisen kuilun, ja väittää, että mentaaliset ominaisuudet redusoituvat fysikaaliseen. Maailmassa on todella siis vain yhdentyypisiä ominaisuuksia, eli fysikaalisia ominaisuuksia. Tämän kannan edustaja voi olla identiteettiteorian kannattaja, jonka mukaan mentaalinen *on* fysikaalista tai eliminativisti, jonka

ratkaisu ongelmaan on vielä yksinkertaisempi: mentaalista ei ole olemassa, joten puhe kuiluista on merkityksetöntä.

Tietoisuuden ja fyysikaalisen suhdetta hämärtää se, että edellistä koskeva tieto on *ensimmäisen persoonan subjektiivista tietoa*, ja jälkimmäistä koskeva *kolmannen persoonan objektiivista tietoa*. Fenomenaalisista faktoista voi saada vain kokemusperäistä subjektista riippuvaista tietoa, kun taas fyysikaalisista faktoista voi saada objektiivista, monien näkökulmien tietoa. Tietoisuuden ja fyysikalismien yhteensovittamisen suurimmat kysymykset kuuluvatkin: *kuinka fyysikaalinen systeemi voi olla tietoinen ja kuinka (subjektiivisen) tietoisuuden voi selittää (objektiivisilla) fyysikaalisilla termeillä*. Määrittelen seuraavaksi vielä lyhyesti muutamia keskeisiä käsitteitä, jotka liittyvät fyysikalistisiin tapoihin yrittää vastata näihin kysymyksiin.

IDENTITEETTITEORIA: Mentaaliset tilat ovat *identtisiä* (fyysikaalisten) aivotilojen kanssa. Tyyppi-identiteettiteorian mukaan jokaisen psykologisen tilan tyyppi on identtinen tietyn fyysikaalisen tilan tyyppin kanssa, esimerkiksi kipu on identtistä C-hermosäikeiden laukeamisen kanssa. Token-identiteettiteorian mukaan esim. eilen kello 13.05 kokemani kivun *esiintymä* on identtinen eilen klo 13.05 tapahtuneen C-hermosäikeiden laukeamisen *esiintymän* kanssa.

FUNKTIONALISMI: Mentaaliset tilat eivät synny aivojen rakenteesta, vaan perustuvat erilaisiin toimintoihin (eli funktioihin) kognitiivisessa systeemissä. Tietoisuus voi rakentua monenlaisten rakenteiden ”päälle”, eikä sen tarvitse olla identtinen jonkin tietyn esim. aivotilan kanssa.

ELIMINATIVISMI: Eliminativismi X:n suhteen kieltää X:n olemassaolon kokonaan. Mielenfilosofian kontekstissa eliminativisti kieltää koko tietoisuuden tai jonkin sen osan (kuten fenomenaalisen tietoisuuden) olemassaolon. Sen ratkaisu tietoisuuden ongelmaan on siis kieltää koko ongelma.

REDUKTIONISMI: Ei kiellä X:n olemassaoloa, vaan väittää että se on ”todellisuudessa” jotain muuta. *Ontologinen reduktio* paljastaa identiteetin, esim. lämpö on molekyylien liikettä. (Mutta voi olla identiteetti ilman reduktiota, esim.

”Aamutähti on iltatähti.”) Mielenfilosofisessa reduktiossa mieli redusoituu ruumiiseen: mentaalinen toiminta redusoituu aivotoimintaan.

SUPERVENIENSSI: Ei voi olla mentaalista eroa ilman fysikaalista eroa. Mentaalinen supervenioituu fysikaalisesta siten, että oliot, jotka ovat samanlaisia fysikaalisilta ominaisuuksiltaan, ovat välttämättä samanlaisia myös mentaalisilta ominaisuuksiltaan. Toisin sanoen supervenienssiteorian mukaan psykofyysinen konditionaali on välttämättä tosi. Donald Davidsonin sanoin:

[S]upervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect (Davidson 1970, 119).

PSYKOFYYSINEN KONDITIONAALI: *Jos* kaikki fysikaaliset totuudet, *niin* kaikki mentaaliset/psykologiset totuudet.

2. Selityksellinen argumentti

The psychophysical gap is presented to us as an interruption in the normal state of affairs: a puzzle, an anomaly, an absence where there ought to be a presence (Van Gulick 2003, 325).

Voiko fysikalistisella selittämistavalla selittää tietoisuuden? Jos hyväksyy fysikalismin, hyväksyy että kaikki mitä on olemassa, on fysikaalista. Mutta hyväksyykö myös, että kaikki on selitettävissä fysikaalisin termein? Asiasta ei vallitse yksimielisyyttä. Eliminativistien (esim. Dennett 1991) mukaan tietoisuus ei ole mitenkään poikkeuksellinen muihin maailman ilmiöihin verrattuna, ja se on täysin selitettävissä fysikaalisin termein. He siis kieltävät koko kuilun. Kun on selittänyt raportoitavuuden, pääsyn omiin sisäisiin tiloihin, informaation käsittelyn ja muut psykologisen pääsy tietoisuuden kysymykset, on selittänyt koko tietoisuuden. Niiden lisäksi ei ole mitään ylimääräistä ”kokemuksen jonkinlaisuutta”, joten ei siis ole minkäänlaista kuiluakaan. Tällainen kanta ei ”ota tietoisuutta vakavasti”, kuten

Chalmers sanoisi. Se hylkää siis kokonaan fenomenalisen tietoisuuden. Mielestäni kuitenkin seuraavat argumentit osoittavat, että tietoisuuteen kuuluu ylimääräinen kysymys kognitiivisten rakenteiden selittämisen jälkeen: miksi se tuntuu jonkinlaiselta? Kuilun kieltäminen ei siis ole älyllisesti rehellinen kanta.

Jotkut fysikalismin kriitikot (esim. David Chalmers, kpl 2.1.) taas muodostavat täysin päinvastaisen päättelyketjun: tietoisuus ei ole selitettävissä fysikaalisin termein, joten tietoisuus ei *ole* fysikaalinen. Hän siis muodostaa episteemisestä kuilusta ontologisen kuilun. Myös keskitien kanta on mahdollinen: vaikka hyväksyisikin, ettei tietoisuus ole *selitettävissä* fysikaalisin termein, voi säilyttää fysikalistisen ontologian (esim. Joseph Levine, kpl 2.2.)

Chalmersin sanoin tietoisuuden suhteen vallitsee *episteeminen asymmetria* (1996, 102). Tieto, että tietoista kokemusta on ylipäättään olemassa, on saavutettavissa pelkästään ensimmäisen persoonan näkökulmasta. Muista ilmiöistä saa tietoa muistakin näkökulmista. Selityksellinen argumentti on nimensä mukaisesti episteeminen: aivojen ja tietoisuuden väliin jäävä kuilu on selityksellinen, toisin sanoen käsitteellinen – ei ontologinen. Se ei ota (välttämättä) kantaa siihen, onko aivojen ja tietoisuuden välissä ”todellista” (ontologista) kuilua, vaan väittää ainoastaan, että osapuolet eivät ole selitettävissä samalla selitysmallilla. Episteeminenkin kuilu kuitenkin osoittaa, että fysikalismi on vaikeuksissa tietoisuuden suhteen.

Though clearly an epistemological matter, I think this intuitive resistance to materialism should bother us a lot (Levine 1983, 356).

Esittelen kolme selityksellisen argumentin muotoa. *What is it like -argumentti* osoittaa, että fenomenalinen ja fysikaalinen ovat luonteeltaan niin erilaisia, ettei jälkimmäisen avulla voi selittää ensimmäistä. *Selityksellisen kuilun argumentti* käsittelee samaa asiaa: fenomenalisen ja fysikaalisen välissä on kuilu (eikä jälkimmäinen siis selitä ensimmäistä). *Selityksellisen argumentin ontologinen muoto*

taas menee askeleen pitemmälle: kuilusta seuraa, että tietoisuus ei *ole* (kaikilta osin) fyysikaalinen.¹⁰

2.1. What is it like -argumentti

Thomas Nagel on esittänyt kuuluisan fysikalismen kritiikkinsä artikkelissaan ”What is it like to be a bat?” (1974). Artikkelin päätarkoituksena on osoittaa, että mentaaliset ilmiöt ovat *subjektiivisia* ilmiöitä ja että tietoisena oleminen tarkoittaa juuri sitä, että on jokin subjektiivinen näkökulma: ”the fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism” (Nagel 1974, 219). Tietoisuus on *näkökulmaisuu*ta, johon pääsee käsiksi vain kulloinkin kyseessä olevasta näkökulmasta. Argumentti on siis selityksellinen: tietoisesta kokemuksesta *jonkinlaisuutta* ei voi selittää fyysikaalisin termein.

Palatkaamme hetkeksi Sokeiden Valtakuntaan. Kuinka Bogota voisi kertoa sokeille millaista on näkeminen? Hänen olisi varmasti mahdollista saada heidät uskomaan, että hänellä on jokin ylimääräinen kyky, sillä hän kykenisi kaikenlaisiin ihmeellisiin suorituksiin, kuten kertomaan kaukana kosketusetäisyydestä kuinka monta sormea koehenkilöllä on pystyssä. Hän voisi myös kertoa heille kuinka näköjärjestelmä toimii. Mutta sitä, *millainen* tämä kyky on, olisi mahdotonta selittää. Se pitäisi kokea itse.

Mikään määrä objektiivista, tieteellistä tietoa ei riitä kertomaan meille myöskään sitä, minkälaista on olla lepakko. Lepakkona olemisen *jonkinlaisuuden* voi tietää vain lepakko itse. Ihminen ei voi päästä tähän subjektiiviseen kokemukseen objektiivisesti käsiksi (eikä tietenkään subjektiivisestikaan – emmehän ole lepakoita), sillä objektiivisuus vaatii juuri riippumattomuutta jostain yhdestä tietystä näkökannasta (subjektiivisuudesta). Koska subjektiivisuus on kuitenkin tietoisuuden olennaisin sisältö, ei objektiivisen kuvauksen (neurologisen tutkimuksen jne.) avulla voi siis periaatteessakaan saada tietoa tietoisuudesta ja tietoisesta kokemuksesta.

¹⁰ Argumenttien esitysjärjestys on myös kronologinen: Nagelin artikkeli on vuodelta 1974, Levinen artikkeli vuodelta 1983 ja Chalmersin artikkelit vuosilta 1995 ja 2002.

Reduktiivinen fysikalismi on siis lähtökohtaisesti mahdottomalta kuulostavan tehtävän edessä: Selittääkseen mentaalisia tiloja sen pitäisi selittää niitä subjektiivisesta näkökulmasta, mutta subjektiivinen näkökulma ei ole reduktiivista fysikalismia. Ilman subjektiivista näkökulmaa se tulee taas selittäneeksi jotain muuta, kuin mistä oli alun perin kysymys. Sivuuttaessaan kokemuksen subjektiivisen luonteen reduktionistiset teorit tulevat sivuuttaneeksi koko tietoisien kokemuksen.

If we acknowledge that a physical theory of mind must account for the subjective character of experience, we must admit that no presently available conception gives us a clue how this could be done. The problem is unique. If mental processes are indeed physical processes, then there is something it is like, intrinsically, to undergo certain physical processes. What it is for such a thing to be the case remains a mystery. (Nagel 1974, 223.)

Kuinka tietoisuutta, joka vaatii ensimmäisen persoonan näkökulman, voisi selittää tieteen kolmannen persoonan näkökulmasta? Nagelin tarkoituksena ei ole itse asiassa osoittaa fysikalismia epätodeksi, vaan vain tuoda esiin, kuinka (ainakaan nykyinen) ymmärryskykymme ei yksinkertaisesti riitä selittämään, kuinka mentaaliset ilmiöt voisivat olla fysikaalisia ilmiöitä. Kyseessä on siis puhdas episteeminen argumentti.

At the present time the status of physicalism is similar to that which the hypothesis that matter is energy would have had if uttered by a pre-Socratic philosopher (Nagel 1974, 224).

What is it like -argumentti tuo esille tietoisuuden ja fysikalismin episteemisen yhteensovittamisen ongelman: fysikaalisen rakenteen ja tietoisuuden välissä on selityksellinen kuilu, ei välttämättä ontologista kuilua. Nagel ei muotoillut fysikalismin kritiikistään samanlaista päättelyketjua kuin modaali- tai tietoargumentti, vaan kritiikki pikemminkin loi pohjan tuleville kirjoittajille. Nagel toi kuitenkin esiin asian, jota fysikalisti ei voi ohittaa: tietoiset tilat ovat subjektiivisia tiloja, eikä niitä voi redusoida molekyyalitasolle samalla tavalla kuin fysiikan ilmiöitä. Nagel olisi voinut lisätä, että ”äskeisestä johtuen fysikalismi on väärässä” – mutta sitä hän ei tehnyt. Lepakkoargumentista olisi voinut muodostaa myös tietoargumentin: Lepakkotutkija Robin tietää kaiken, minkä voi kirjoista oppia

lepakoiden kaikuluotainsysteemistä. Hän tietää täsmälleen, mitkä hermosolut aktivoituvat milläkin kaikuluotaamisen hetkellä. Eräänä aamuna hän herää ja huomaa saaneensa itsekin kaikuluotaamisen kyvyn. Oppiiko hän jotain uutta kaikuluodatessaan ensimmäisen kerran? Nagelin artikkelin sisällön kanssa olisi myös täysin sopusoinnussa zombiargumentti: Aivotutkimus ei riitä kertomaan, kokeeko lepakko lepakkouttaan minkäänlaisena, vaikka sen fysikaalinen rakenne olisikin täsmälleen sama kuin muiden lepakoiden.¹¹ Siispä reduktionismi on väärässä.

Nagel itse ei kuitenkaan *argumentoi* sen puolesta, että fysikalismi olisi väärässä. Hänen esiintuomansa reduktiivisen fysikalismin ongelmallisuus on kuitenkin ollut tärkeänä innoittajana muille fysikalismia vastaan suunnatuille argumenteille, kuten tietoargumentille.

2.2. Selityksellinen kuilu

Selityksellisen kuilun (*explanatory gap*) argumentissa tietoisuuden ja aivotilojen väliin jäävä kuilu on – nimensä mukaisesti – vain selityksellinen. Termiä on alun perin käyttänyt Joseph Levine artikkelissaan ”Materialism and Qualia: The Explanatory Gap” (1983). Levinen tarkoituksena on osoittaa, että tietyt psykofyysiset identiteettiväitteet (mentaalinen tila *m* on yhtä kuin fyysinen tila *f*) jättävät sisälleen merkittävän selityksellisen kuilun – ja se on ongelmallista fysikalismille.

Levine pyrkii muuttamaan Kripken (1980)¹² metafyyssisen väitteen ”*jos psykofyysiset identiteettiväitteet ovat kuviteltavissa vääriksi, ne myös ovat väärinä*” epistemologiseksi väitteeksi. Levinen mukaan Kripke on oikeilla jäljillä luottaessaan tietynlaiseen materialismin vastaiseen *intuitioon* tietoisesta kokemuksesta suhteen, mutta hän ei usko Kripken tapaan, että siitä seuraa, että psykofyysisten identiteettiväitteiden täytyy olla epätosia.

Rather, I think it supports a closely related epistemological thesis – namely, that psycho-physical identity statements leave a significant *explanatory gap*, and, as a corollary, that

¹¹ Tässä tulee esiin myös *muiden mielten ongelma*.

¹² Käsittelem Kripken argumenttia kappaleessa 3.1.

we don't have any way of determining exactly which psycho-physical identity statements are true. One cannot conclude from my version of the argument that materialism is false, which makes my version a weaker attack than Kripke's. Nevertheless, it does, if correct, constitute a problem for materialism[.] (Levine 1983, 354.)

Levine (1983, 354) käyttää esimerkkinä (Kripkeltä tuttuja) väitelauseita, jotka sisältävät identiteettiteorian kannattajan käsityksiä ilmiöiden redusoitumisesta alemmalle tasolle.

(1) Kipu on C-hermosäikeiden laukeamista.

(2) Lämpö on molekyylien liikettä.¹³

Kripken referenssiteorian mukaan ”lämpö” ja ”molekyylien liike”, kuten myös ”kipu” ja ”C-hermosäikeiden laukeaminen” ovat *kiinteitä nimittäjiä*,¹⁴ joten jos lauseet (eli niiden sisältämät identiteettiväitteet) ovat ylipäättään tosia, ne ovat *välttämättä* tosia. Lauseet ovat kuitenkin kuviteltavissa epätosiksi, mikä on ongelmallista, jos ne kerran ovat välttämättä tosia. Jos kuvittelee kipua ilman C-hermosäikeiden laukeamista tai lämpöä ilman molekyylien liikettä, kuvittelee loogisesti mahdotonta maailmaa. Lauseen (2) näennäisen kontingenssin voi kuitenkin sulkea pois, mutta lauseen (1) ei. Lauseen (2) prima facie -kontingenssi johtuu sekaannuksesta. Todella kontingentti lause olisi seuraava:

(2') Ilmiö, jonka koemme lämpöaistimuksessa, ja joka aiheuttaa elohopean laajenemisen lämpömittareissa jne. on molekyylien liike.

Ero on siinä, että lauseen (1) kohdalla ei voi erottaa sitä, miten ilmiö (kipu) ilmenee meille ja itse ilmiötä (kivun aistimusta). Lause (1) on siis todella kontingentti. Lämpö

¹³ (1) *Pain is the firing of C-fibers.* (2) *Heat is the motion of molecules.* (Levine 1983, 354.) Näitä samoja esimerkkejä on käytetty vuosikymmeniä alan kirjallisuudessa, vaikka jo Kripke huomautti (1980, 149), ettei hän oikeastaan tiedä C-hermosäikeistä mitään (paitsi että niiden laukeamisen sanotaan liittyvän kipuun). Tällä ei kuitenkaan ole merkitystä, C-hermosäikeiden laukeamisen voi korvata halutessaan muulla fysikaalisella ilmiöllä argumentin muuttumatta.

¹⁴ (Engl. *rigid designator.*) Termi on kiinteä nimittäjä, kun se viittaa (nimittää, tarkoittaa) samaan kohteeseen kaikissa mahdollisissa maailmoissa, missä tuo asia esiintyy, mutta ei mihinkään muuhun. Esimerkiksi ”Suomen presidentti vuonna 2011” ei ole kiinteä nimittäjä, koska presidentti olisi voinut olla kuka tahansa (jossain mahdollisessa maailmassa). Sen sijaan kukaan muu ei olisi voinut olla ”Tarja Halonen” kuin Tarja Halonen, eli kyseessä on kiinteä nimittäjä (Kripke 1980, 48.) Palaan Kripken teoriaan tarkemmin kappaleessa 3.1.

ja lämmön aistimus sen sijaan on mahdollista erottaa toisistaan. Tämä osoittaa, että psykofyysiset identiteettiväitteet ja muut identiteettiväitteet toimivat eri tavalla, ja tämä ero tekee Levinen (1983, 355) mielestä ensin mainitut epäuskottaviksi.

Lause (1) jättää selityksellisen kuilun, lause (2) ei. Lause (2) ei jätä mitään selityksen ulkopuolelle, sillä kemian ja fysiikan teoriat selittävät sen tyhjentävästi. Fysikalistinen teoria selittää, kuinka lämpö redusoituu molekyylien liikkeeseen: kunhan siis tunnemme tarpeeksi fysiikan ja kemian teorioita, ymmärrämme miten molekyylien liike liittyy kausaalisesti lämpöön. Sen sijaan kvalioiden fysikaaliseen palautumista kemian ja fysiikan teoriat eivät selitä. Mikään määrä tieteellistä tietoa ei selitä, mikä yhteys C-hermosäikeillä ja kivulla on (*mutatis mutandis* muiden fenomenaalisten ilmiöiden suhteen). Lauseen (1) voi kyllä selittää kausaalisella tavalla: terävän metalliesineen työntäminen ihoon vaikuttaa tiettyihin hermopäihin, jotka vaikuttavat C-hermosäikeisiin, jotka vaikuttavat hermostoa pitkin lopulta aivoihin. Äskeinen on varmasti totta, ja selittää tärkeän (voisi sanoa psykologispääsytietoisuuden) osan kivun käsitteestä. Mutta kivun käsitteeseen kuuluu muutakin kuin sen kausaalinen rooli – tarkemmin sanottuna sen fenomenaalinen luonne, eli *miltä* se tuntuu. C-hermosäikeiden laukeaminen ei lainkaan selitä, miksi kipu tuntuu juuri sellaiselta kuin se tuntuu, vaan selityksessä on täydellinen aukko, *explanatory gap*. Yhteys on täysin mystinen.

Koska ei ole mitään, mikä selittäisi C-hermosäikeiden laukeamisen ja kivun yhteyden, tulee mahdolliseksi *kuvitella* C-hermosäikeiden laukeaminen ilman kipua, sekä kipu ilman C-hermosäikeiden laukeamista.

So long as the two states of affairs seem arbitrarily stuck together in this way, imagination will pry them apart. Thus it is the non-intelligibility of the connection between the feeling of pain and its physical correlate that underlies the apparent contingency of that connection. (Levine 1983, 359.)

Yleensä tällaiset Kripke-tyyppiset argumentit on suunnattu identiteettiteoriaa vastaan, mutta Levine (1983, 355–356) pyrkii laajentamaan argumentin koskemaan myös funktionalismia. Kivun funktionaalinen selitys voi olla monenlainen (toteutua monilla eri tavoilla) – vaikka emme tiedä millainen se tarkalleen on, tiedämme

kuitenkin millaisessa muodossa se on. Voidaan siis olettaa kivun täydellinen funktionaalinen selitys F. Funktionalismi väittää, että:

(3) Kipu on tilassa F olemista.

Levinen mukaan Kripken teoriaa voi soveltaa tähän lauseeseen samalla tavalla kuin lauseisiin (1) ja (2). Voi kuitenkin kysyä, onko F todella kiinteä nimittäjä. F ei viittaa samaan kohteeseen kaikissa mahdollisissa maailmoissa, sillä se voi toteutua useilla tavoilla. Toisaalta kyseessä on kuitenkin aina sama kivun funktionaalinen selitys. Jos F ei ole kiinteä nimittäjä, ei Levinen yritys laajentaa argumenttia myös funktionalismiin onnistu. Kuten Levine (1983, 356) huomauttaa, funktionalismi on joka tapauksessa ongelmallinen kanta, sillä funktionaaliset kuvaukset ovat niin abstrakteja, että niihin sisältyy *liikaa* mahdollisuuksia. Myöhemmin esittelemäni (kappaleessa 3) puuttuvan ja käänteisen kvalian argumentit osoittavat, että funktionaaliset selitykset ovat abstraktisuutensa takia avoimia useille mahdollisuuksille, eivätkä ne kykene tämän takia vangitsemaan kvalioita. Jotkut funktionalismin kannattajat¹⁵ ovatkin tämän takia siirtyneet takaisin identiteettiteoreetikoiden leiriin.

Selityksellisen kuilun argumentista seuraa, että vaikka jotkin psykofyysiset identiteettiväitteet voivatkin olla tosia, *meillä ei ole mitään keinoa tietää, mitkä niistä ovat tosia*. Tämä ei ole kovin suotava ominaisuus fysikalistiselle teorialle. Toinen argumentin esiintuoma fysikalismin ongelma on samanlaisuuden/erilaisuuden asteen määrittelemättömyys. Kun oletetaan, että kipu on identtinen tietyn fysikaalisen tilan (C-hermosäikeiden laukeamisen) kanssa, tämä fysikaalinen tila on päätelty *ihmisruumiin* toiminnasta. Entä jos satumme tapaamaan avaruusolioita, jotka vaikuttavat käytöksensä ja funktioidensa perusteella kokevan fenomenaaliset kokemukset (kuten kivun) juuri samalla tavalla kuin me? Tuntevatko ne kipua? Jos kipu kerran on C-hermosäikeiden laukeamista, ja avaruusolioilla ei ole koko C-hermosäikeitä, ne eivät siis voi tuntea kipua. Mutta vaikka tämän (sinänsä epäilyttävältä kuulostavan) seurauksen hyväksyisikin, se tuo esiin ongelman, jota identiteettiteoreetikokkaan ei voi ohittaa: kuinka voimme päättää, milloin olio on

¹⁵ Esim. Shoemaker (1982).

tarpeeksi samanlainen tai liian erilainen? (Levine 1983, 359–360.) Tuntevatko koirat kipua? Entä simpukat?

Viimeksi mainittu ongelma tulee esiin myös *moninaisen toteutuvuuden argumentissa*, jota käsittelemme kappaleessa 3.5. Mutta toisin kuin moninaisen toteutuvuuden argumentti, Levine pyrkii kritisoimaan (identiteettiteorian lisäksi) myös funktionalismia. Hänen mukaansa (1983, 360) funktionalismiin toimii sama kritiikki kuin identiteettiteoriaankin: jos väitämme, että kipu on olemista jossain tietyssä funktionaalissa tilassa, millä tavalla voimme mitata, onko olio tarpeeksi samanlaisessa/erilaisessa funktionaalissa tilassa? (Ja tästä seuraa – taas – puuttuvan- ja käänteisen kvalian ongelma.)

[W]ithout this explanatory gap filled in, facts about the kind or existence of phenomenal experiences of pain in creatures physically (or functionally) different from us become impossible to determine. This, in turn, entails that the truth or falsity of (1), while perhaps metaphysically factual, is nevertheless epistemologically inaccessible. This seems to be a very undesirable consequence of materialism. (Levine 1983, 360.)

Levinen muotoilemana selityksellinen kuilu ei koske ontologiaa, vaan epistemologiaa. Lauseen ”kipu on C-hermosäikeiden laukeamista” totuusarvo voi siis olla ontologisesti tosi tai epätosi, saavuttamattomissa se on vain epistemologisesti. Levine ei juurikaan perustele, miksi epistemologisesta argumentista ei seuraa ontologinen argumentti. (Paitsi että oletus on kätevä, koska tällä tavoin ei tarvitse luopua materialistisesta ontologiasta.) Sama pätee toisinkin päin: Kripke ja Chalmerskaan eivät selitä tyhjentävästi, miksi epistemologisesta seuraa ontologinen. Onkin siis todella hyvä kysymys, mikä epistemologisen ja ontologisen ero tässä yhteydessä oikein on – jos mikään. Tästä kysymyksestä riippuu mielestäni koko fysikalismin selviytymismahdollisuus.

2.3. Selityksellisen argumentin ontologinen muoto

Myös Chalmersin muotoilema selityksellinen argumentti perustuu aukkoon fysikaalisen ja fenomenaalisen selityksen välillä. Chalmersin termein (1995; ks. kappale 1.1.) tietoisuuteen kuuluu *helppoja ongelmia* (psykologisen pääsy tietoisuuden kysymykset) sekä *vaikea ongelma* (tietoisuuden fenomenaalinen puoli). Nämä kaksi tietoisuuden osaa eivät ole samalla tavalla selitettävissä: helpot ongelmat voi selittää selittämällä niiden fysikaaliset (funktionaaliset) piirteet, mutta Chalmersin mukaan mikään määrä fysikaalista tietoa ei voi selittää tietoisuuden vaikeaa ongelmaa – tietoista kokemusta, sitä että kokeminen on *jonkinlaista*.

One can argue that by the character of physical explanation, physical accounts explain *only* structure and function, where the relevant structures are spatiotemporal structures, and the relevant functions are causal roles in the production of system's behavior. And one can argue as above that explaining structures and functions does not suffice to explain consciousness. If so, no physical account can explain consciousness. (Chalmers 2002a, 248.)

Psykologisen pääsy tietoisuuden (helpot) ongelmat voi kyllä selittää tieteellisesti, siis fysikaalisin termein. Fenomenaalinen tietoisuus (vaikea ongelma) on luonteeltaan erilainen: jos yrittää selittää sen rakenteet ja funktiot, tuleekin selittäneeksi vain siihen liittyvien helppojen ongelmien rakenteita ja funktioita. Itse vaikea ongelma, se että *tietoisuus on jonkinlaista*, jää selittämättömäksi. Chalmersin (2002a, 248) argumentti alkaa seuraavanlaisesta päättelystä:

- (1) Fysikalistisella selitystavalla voi selittää vain rakenteita ja funktioita.**
- (2) Rakenteiden ja funktioiden selittäminen ei riitä tietoisuuden selittämiseen.**
- (3) Fysikalistinen selitysmalli ei voi selittää tietoisuutta.**

Chalmersin mukaan (2002a, 253–254) tietoisuuden poikkeuksellisesta luonteesta (muihin maailman ilmiöihin verrattuna) seuraa, että tietoisuuden ja fysikaalisen välinen episteeminen kuilu on erilainen muiden alojen episteemisiin kuiluihin verrattuna. Selittääkseen geenit täytyy vain selittää mekanismit, jotka varastoivat ja siirtävät perinnöllisyysinformaatiota sukupolvelta toiselle. Selviää, että DNA

toteuttaa tämän mekanismin, ja kun on selvitetty kuinka se toimii, on selitetty geenit. Toisin sanoen: jos tietää kaiken perinnöllisyystieteestä, voi siitä johtaa deduktiivisesti geenien toiminnan. Vaikuttaa myös siltä, että emme voi kuvitella fysikaalisesti identtistä maailmaa, jossa ei olisi geenejä. Siispä fysikaalisen selityksen (maailmasta) ja geenien selityksen välissä ei ole samanlaista episteemistä kuilua kuin fysikaalisen selityksen (maailmasta) ja tietoisuuden selityksen välissä. Geenien ja DNA:n välinen suhde ei siis ole *episteemisesti primitiivinen*, vaan identiteetin voi johtaa deduktiivisesti (täydellisestä) fysikaalisesta selityksestä. Selittäminen onnistuu vastaavasti muiden fysikaalisten ilmiöiden sekä psykologisen pääsytietoisuuden kysymyksen suhteen.

We can put this by saying that while the identity between genes and DNA is empirical, it is not *epistemically primitive*: the identity is itself deducible from the complete physical truth about the world (Chalmers 2002a, 254).

Tietoisuus ei Chalmersin mukaan ole selitettävissä samalla tavalla. Tietoiseen kokemukseen liittyvät faktat eivät voi olla suora seuraus fysikaalisesta selityksestä, sillä on käsitteellisesti mahdollista, että fysikaaliset prosessit tapahtuisivat ilman kokemusta (Chalmers 1995, 8). Vaikka selittäisi kaikki käyttäytymiseen ja kognitiiviseen toimintaan (psykologiseen pääsytietoisuuteen) liittyvät kysymykset, jäisi kuitenkin yksi kysymys vaille vastausta: miksi niihin liittyy tietoinen kokemus?

By contrast, [...] the identification between consciousness and physical or functional states is epistemically primitive: the identity is not deducible from the complete physical truth (Chalmers 2002a, 254).

Chalmersin mukaan siis siitä, että fysikalistinen selitysmalli ei voi selittää tietoisuutta, seuraa, että tietoisuuden ja fysikaalisen suhde on jonkinlainen primitiivinen periaate. Argumenttiin tulee siis vielä yksi premissi, joka johtaa uuteen johtopäätökseen:

(4) Se, mikä ei ole selitettävissä fysikaalisin termein, ei ole fysikaalista.

(5) Siispä fysikalismi (tietoisuuden suhteen) on väärässä.

Chalmers siis johtaa episteemisestä argumentista ontologisen argumentin. Hänen mukaansa primitiivisiä periaatteita tulee kohdella *fundamentaalisina lakeina*. Fundamentaalisia entiteettejä ei voi selittää minkään perustavammanlaatuisen avulla, ne eivät enää redusoidu mihinkään. Fysiikassa tällaisia ilmiöitä ovat esimerkiksi massa ja aika-avaruus. Niitä ei voi selittää minkään yksinkertaisemman avulla. (Mutta fundamentaalisuus ei kuitenkaan estä muotoilemasta tällaisista ilmiöistä teorioita.) Chalmersin mukaan tietoinen kokemus tulee lisätä fundamentaalisten lakien joukkoon:

I suggest that a theory of consciousness should take experience as fundamental. We know that a theory of consciousness requires the addition of *something* fundamental to our ontology, as everything in physical theory is compatible with the absence of consciousness. We might add some entirely new nonphysical feature, from which experience can be derived, but it is hard to see what such a feature would be like. More likely, we will take experience itself as a fundamental feature of the world, alongside mass, charge and space-time. (Chalmers 1995, 9).

Chalmers huomauttaa (2002a, 254), että selityksellisen argumentin kannattajista ne, jotka hyväksyvät sen vain kohtaan 3 asti, tulevat kannattaneeksi, että tietoisuus on primitiivinen periaate, mutta se kuitenkin redusoituu johonkin muuhun (fysikaaliseen). Kanta on Chalmersin mielestä ristiriitainen, sillä hänen mukaansa primitiivisiä periaatteita tulee kohdella fundamentaalisina lakeina – eivätkä fundamentaaliset lait redusoidu mihinkään. Chalmersin mielestä he kannattavat tällaista näkemystä lähinnä vain voidakseen säilyttää fysikalistisen ontologian.

Chalmers (1995, 10; 2002a, 254) itse pitää tietoisuuden ja fysikaalisen suhdetta fundamentaalisenä luonnonlakina, ja näin ollen kannattaa (non-reduktiivista) ominaisuusdualismia. Hänen mukaansa tällaisessa mallissa ei ole mitään ristiriitaista fysiikan kanssa, sillä se ei lainkaan puutu fysiikan lakeihin, vaan on vain ylimääräinen lisäys niiden ohella. Hän kutsuukin kantaansa *naturalistiseksi dualismiksi*, sillä se on täysin yhteensopiva kaikkien olemassa olevien fysiikan lakien kanssa.

2.4. Yhteenveto selityksellisestä argumentista

Selityksellisestä kuilusta voi siis päätyä monenlaisiin lopputuloksiin. Nagel tuo vain esiin, kuinka fenomenaalinen ja fysikaalinen tieto ovat luonteiltaan erilaisia, mutta ei väitä, etteikö fysikaalinen *periaatteessa* voisi selittää fenomenaalista.¹⁶ Levine taas pitää kuilua selvänä osoituksena, ettei fysikaalinen voi selittää fenomenaalista. Levinen argumentti on kuitenkin puhtaasti epistemologinen; vaikka tietoisuutta ei voi selittää fysikaalisilla termeillä, ei siitä kuitenkaan seuraa, etteikö se olisi fysikaalinen. Chalmersin mielestä Levinen ratkaisu ei ole johdonmukainen: jos tietoisuutta ei voi selittää fysikaalisin termein, se ei myöskään ole fysikaalinen. Sen sijaan tietoinen kokemus tulee käsittää redusoitumattomana fundamentaalisenä luonnonlakina. Kuilun ylitsepääsemättömyyteen voi siis suhtautua eriasteisesti: kuilu voi olla este vain nykyisen käsityskykymme puutteiden takia (Nagel), ihmisymmärryksen luonteen rajallisuuden takia (Levine), tai se voi olla ihmisymmärryksen lisäksi kaikenlaisen muunkin selittämisen ulkopuolella (Chalmers).

Mielestäni selityksellinen argumentti osoittaa, että tietoisuuden ja fysikaalisen välissä todella on selityksellinen kuilu. Kuilun kieltäjät (esim. Dennett 1991) eivät tule puhuneeksi samasta asiasta väittäessään, että ”kuilu” on selitetty, kunhan psykologisen pääsy tietoisuuden kysymykset on selitetty. Kuilun hyväksymisen jälkeen siitä kuitenkin edetään usein johtopäätöksiin liian kevein perustein. Päätelmissä on syytä huomioida, että siitä, että fysikaalinen ei selitä tietoisuutta, ei kuitenkaan voi suoraan päätellä, ettei fysikaalinen *voisi* selittää tietoisuutta. Ja siitä, ettemme tällä hetkellä ymmärrä kuinka selitys olisi mahdollinen, ei seuraa, että selitys todella olisi mahdoton. Nagelin ja Levinen argumentit ovatkin siis uskottavampia kuin Chalmersin ontologiseen lopputulokseen päätyvä argumentti.

Chalmersin päättelyn ongelma on siinä, että silta episteemisestä ontologiseen on jossain määrin kyseenalainen. On varmasti totta, että fundamentaaliset lait eivät redusoidu mihinkään (sillä redusoitumattomuus on sisäänrakennettuna

¹⁶ Hän spekuloi jonkinlaisen *objektiivisen fenomenologian* mahdollisuudella, jossa subjektiiviset kokemukset olisivat selitettävissä kolmannen persoonan näkökulmasta (Nagel 1974, 224). Ehdotus kuulostaa ristiriitaiselta kaiken sen kanssa, mitä Nagel on osoittanut subjektiivisen ja objektiivisen kokemuksen erosta. Tuon tämän kuitenkin esiin esimerkinä kannasta, joka pitää kuilua periaatteessa ylitettävänä.

fundamentaalisuuden käsitteessä). On kuitenkin kyseenalaista, onko tietoisuuden ja fyysikaalisen suhde Chalmersin kuvaama primitiivinen yhteys. Chalmersin argumenttiin kuuluu hyvin vahva oletus, että kuilu *ei ole periaatteessakaan* ylitettävissä. Tästä oletuksesta seuraa melko helposti, että koska kuilu on niin jyrkkä episteeminen ero, niin sen täytyy olla myös metafyyssinen ero. Onkin siis syytä tarkastella Chalmersin argumentin ensimmäistä johtopäätöstä (3. Fysikalistinen selitysmalli ei voi selittää tietoisuutta). Onko se todella tosi, vai olisiko oikeutettu johtopäätös kuitenkin vain ”(Nykyinen) fysikalistinen selitysmalli ei selitä tietoisuutta”? Jos jälkimmäinen onkin oikea johtopäätös, niin Chalmersin ontologinen argumentti kaatuu.

Onko selityksellinen kuilu sitten ongelma fysikalismille?¹⁷ Jos kuilu on (vain) episteeminen, ei ontologinen fysikalismi ole vaarassa. Tässä mielessä kuilu ei siis varsinaisesti uhkaa fysikalismia. Mutta episteeminenkin kuilu kuitenkin osoittaa, että koska fysikalismi ei selitä tietoisuutta, niin fysikalismissa täytyy olla jotain vialla.

3. Modaaliargumentti

Modaaliargumentti¹⁸ fysikalismia vastaan perustuu ajatukseen siitä, että fyysikaaliset ilmiöt *ovat kuviteltavissa* ilman mentaalisia ilmiöitä. Mentaalisten ja fyysikaalisten ilmiöiden erillisyyden *kuviteltavuus*¹⁹ perustuu siihen, että on mahdollista kuvitella maailma, joka on fyysikaalisesti samanlainen kuin meidän maailmamme, mutta josta puuttuu tietoinen kokemus (*puuttuva kvalia*). Samoin on myös kuviteltavissa fyysikaalisesti identtinen maailma, jossa tietoiset kokemukset ovat *erilaisia* kuin

¹⁷ On mahdollista kannattaa näkemystä, että kuilu jopa vahvistaa fysikalismin oikeutusta (esim. Van Gulick 2003).

¹⁸ Kutsun modaaliargumenttia tässä myös *kuviteltavuus*-argumentiksi (*the conceivability argument*), joka on tarkalleen ottaen osuvampi nimitys, sillä argumentti alkaa kuviteltavuudesta (eli *kuviteltavissaolevuudesta*), josta on tarkoitus edetä (modaaliseen) mahdollisuuteen. Englannin termi *'conceivability'* (ajateltavissa/kuviteltavissaolevuus) eroaa selvästi termistä *'imaginability'* (kuviteltavissaolevuus). Suomenkielisissä ilmauksissa (*ajateltavissa-* ja *kuviteltavissaolevuus*) ei ole näin suurta eroa, joten käytän jälkimmäistä silläkin uhalla, että se herättää konnotaatiota mielikuvituksellisesta kuvittelemisesta. Pidän sitä kuitenkin *ajateltavissaolevuutta* parempana, koska kyse on kuitenkin juuri *kuvittelusta* eikä *ajattelusta*.

¹⁹ *Kuviteltavuus* ei ole kovin tarkka käsite, sitä on kuvailtu mm. seuraavilla tavoilla: ”rationaalinen intuitio, että on mahdollista, että P”, ”huomio, että ei-P ei ole välttämätön”, ”kuvittelu, että P”, ”oletus, että P”, ”ymmärrys propositiosta, että P”, ”kuvailu tilanteesta, jossa P vallitsee”, ”koherentti kertomus, jossa P vallitsee” jne. (Gendler & Hawthorne 2002, 7).

meidän maailmassamme (*käänteinen kvalia*). Argumentti toimii toisinkin päin: on mahdollista kuvitella *erilainen* fysikaalinen maailma, jossa on *samanlainen* tietoinen sisältö (*moninainen toteutuvuus*). Argumentin kannattajien mukaan siitä, että tällainen kuvittelemine on loogisesti mahdollista,²⁰ seuraa että se on metafysisesti mahdollista, joten fysikalismi on väärässä.

Argumentin taustalla vaikuttaa Descartesin ajatus, jonka mukaan mieli ja ruumis ovat toisistaan erillisiä substansseja, koska mieli voi olla olemassa ilman ruumista. Kuten Kripke huomauttaa (1980, 145) Descartes olisi voinut päätellä saman toisinkin päin: koska ruumis voi olla olemassa ilman mieltä, täytyy niiden olla erillisiä substansseja. (Ja on selvää, että ruumis voi olla olemassa ilman mieltä – kuolleen ruumiina.)

Seuraavaksi tutkin tarkkaavaisena, mikä olin, ja huomasin voivani kuvitella, ettei minulla ollut ruumista ja ettei ollut maailmaa eikä mitään paikkaa, missä olin. En kuitenkaan voinut kuvitella, etten itse olisi olemassa, vaan päinvastoin juuri siitä, että ajattelin epäileväni muiden asioiden totuutta, seurasi aivan ilmeisesti ja aivan varmasti, että olin olemassa. [...] Näin ollen tämä *minä* – toisin sanoen sielu, jonka vuoksi olen mitä olen – on kokonaan ruumiista erillinen. (Descartes 2001, 140.)

Modaaliargumentti kyseenalaistaa fysikalistisen käsityksen a priori -supervenienssista – siitä, että tietoisuus supervenioituu loogisesti fysikaalisesta. Argumentti perustuu siihen, että on mahdollista kuvitella, että näin ei tapahdu. Fysikaalisen ja tietoisuuden suhteen voi kuitenkin käsittää myös aposterioriseksi välttämättömyydeksi: väitettä onkin usein käytetty modaaliargumentin kritiikkinä (ks. kappale 3.4.) Jaan modaaliargumentin neljään osaan: 1. *Kripken teoriaan*, 2. *Puuttuvan kvalian*, 3. *Käänteisen kvalian* ja 4. *Moninaisen toteutuvuuden* argumentteihin. Puuttuvan kvalian ja käänteisen kvalian argumentit vastustavat ennen kaikkea supervenienssiteoriaa ja funktionalismia, sillä niiden mukaan on kuviteltavissa, että samat *funktionaaliset toiminnot* voivat tuottaa erilaisia mentaalisiä sisältöjä (tai olla tuottamatta niitä lainkaan). Kripken teoria ja moninaisen toteutuvuuden argumentti taas vastustavat reduktiivista fysikalismia ja identiteettiteoriaa (jälkimmäinen on alun perin muotoiltu suorastaan puolustamaan

²⁰ Käytän ilmausta *looginen mahdollisuus*, mutta en viittaa sillä niinkään logiikkaan, vaan pikemminkin kyseessä on *käsitteellinen* tai *episteeminen mahdollisuus*.

funktionalismia). Modaaliargumentti johtaa myös lisäkysymykseen (jota käsittelemme kohdassa 3.4): seuraako jonkin (loogisesta) *kuviteltavuudesta* sen (metafyysinen) *mahdollisuus*?

3.1. Kripken psykofyysisten identiteettien kritiikki

Kripke esitti fysikalistisen identiteettiteorian kritiikkensä 1970-luvun alussa.²¹ Kritiikin vaikutus näkyy kaikissa modaaliargumentin muodoissa, joten esittelen sen tässä pääpiirteittäin. Kripken (1980, 148) mukaan ”lämmön” ja ”molekyyliden liikkeen” sekä ”kivun” ja ”C-hermosäikeiden laukeamisen” välistä identiteettiä pidetään yleensä kontingenttina. Kripken referenssiteorian mukaan kuitenkin kaikki identiteettiväitteet, jotka koostuvat *kiinteistä nimittäjistä ovat välttämättä tosia, jos ovat tosia ylipäätään*. ”Lämpö” ja ”molekyyliden liike” ovat kiinteitä nimittäjiä, joten niiden välinen identtisyys on siis välttämätön, jos se on tosi. Vaikka voisi kuvitella, että voisi olla mahdollinen maailma, jossa lämpö olisi jotain muuta kuin molekyyliden liikettä, tämä ei pidä paikkaansa. Kyseessä ei nimittäin silloin olisi lämpö, vaan vain jokin lämpöä muistuttava asia. Lämpö sen sijaan on aina molekyyliden liikettä.

Sama välttämättömyys pätee myös kipuun ja C-hermosäikeiden laukeamiseen (jos identtisyys on tosi).²² On mahdollista kuvitella maailma, jossa kipua esiintyisi ilman C-hermosäikeiden laukeamista ja toisaalta maailma, jossa esiintyisi C-hermosäikeiden laukeamista ilman kipua. Voisiko kyseessä siis olla näennäinen kipu, kuten edellisessä esimerkissä oli näennäinen lämpö? Voiko kivun erottaa kipumaisesta tuntemuksesta?

’[P]ain’ is a rigid designator of the type, or phenomenon, it designates: if something is a pain it is essentially so, and it seems absurd to suppose that pain could have been some phenomenon other than the one it is. The same holds for the term ‘C-fiber stimulation’, provided that ‘C-fibers’ is a rigid designator, as I will suppose here. [...] Thus the identity of

²¹ Hän esitti sen kuuluisassa luentosarjassaan *Naming and Necessity*, samanniminen kirja julkaistiin myöhemmin (Kripke 1980).

²² Kipu on klassinen esimerkki ilmiöstä, joka on olemassa vain kun joku (subjekti) kokee sen. Se vaatii ruumiin ja tiedostetuksi tuleminen. Esimerkkinä olisi tietenkin voinut olla mikä tahansa muukin aistimus ja sitä vastaava aivotila -pari.

pain with the stimulation of C-fibers, if true, must be necessary. (Kripke 1980, 148–149.)

Joko kipu on siis aina (kaikissa mahdollisissa maailmoissa) välttämättä sama asia kuin C-hermosäikeiden laukeaminen, tai sitten nämä kaksi ilmiötä eivät ole identtisiä. Identiteettiteorian kannattajan täytyy siis sitoutua kantaan, jonka mukaan ei voi olla C-hermosäikeiden laukeamista, joka ei ole kipua, eikä myöskään kipua, joka ei ole C-hermosäikeiden laukeamista. Kripken (1980, 149) mukaan tämä kuulostaa epäintuitiiviselta,²³ sillä molemmat vaihtoehdot ovat kuviteltavissa. Kripken ratkaisu on siis, että kivun ja hermosäikeiden välinen identtisyys ei ole tosi.

Identiteettiteorian kannattaja voi yrittää puolustautua väittämällä, että kivun ja C-hermosäikeiden suhteen kontingenssi on samalla tavalla näennäistä kuin lämmön ja molekyylien liikkeen tapauksessa. Lämmön suhteen näennäinen kontingenssi on selitettävissä pois, sillä on mahdollista erottaa lämpö ja *lämmöntuntemus* toisistaan. Jos joku siis väittää, että lämpö voisi olla jotain muuta kuin molekyylien liikettä, hän väittää oikeastaan, että joku olisi voinut aistia jonkin ilmiön samalla tavalla kuin lämmön – siis saamalla lämpöaistimuksen. Hän tarkoittaa siis, että voisi olla olioita, jotka eivät saa lämpöaistimusta molekyylien liikkeestä, vaan sen sijaan jostain muusta. Tällaiset oliot voisivat olla, *kvalitatiivisesti*, samassa episteemisessä tilanteessa kanssamme: he voisivat käyttää kiinteää nimittäjää ilmiöstä, joka tuottaa lämmöntuntemuksen heissä. Tuntemusta ei aiheuttaisi kuitenkaan molekyylien liike, joten kyseessä ei olisi lämpö. (Kripke 1980, 150–151.)

Jos kivun ja hermosäikeiden välisen kontingenssin voisi osoittaa näennäiseksi vastaavalla tavalla, niin identiteettiteoria pelastuisi. Mutta argumentti ei toimi analogisesti kivun ja C-hermosäikeiden laukeamisen tapauksessa, sillä niiden välillä *ei voi* olla samanlaista näennäistä kontingenssia kuin lämmön ja molekyylien liikkeen välisessä suhteessa. Kontingenssi kuitenkin on, ja koska se ei voi olla näennäistä, sen täytyy olla aitoa. Se johtuu fenomenaalisten ominaisuuksien palautumattomuudesta fysikaaliseen: lämmön voi erottaa ”lämmön aistimisesta”,

²³ Kripke itse ei tarkenna, miksi seuraus on epäintuitiivinen. Moninaisen toteutuvuuden argumentti käsittelee juuri tätä piirrettä: sen mukaan on terveen järjen mukaista ajatella, että ihmisten lisäksi muutkin oliot voivat tuntea kipua, vaikka niillä ei tapahtuisikaan C-hermosäikeiden laukeamista. Tämän vastustaminen on kieltämättä epäintuitiivista. Toisin päin – C-hermosäikeet ilman kipua, eli zombi – on kyllä kuviteltavissa, mutta en tiedä kuinka intuitiivisesti.

mutta kipua ei voi erottaa ”kivun aistimisesta” millään tavalla. *Kipumainen tuntemus* ei voi olla mitään muuta kuin kipua. Lämmön tapauksessa on mahdollista kuvitella, että on olemassa molekyylien liikettä ilman *lämmöntuntemusta*, mutta koska kipua ei voi erottaa *kivuntuntemuksesta*, niin analogisen esimerkin saadakseen tulee esittäneeksi, että voi olla C-hermosäikeiden laukeamista ilman kivuntuntemusta, siis *kipua*.

Thus pain, unlike heat, is not only rigidly designated by ‘pain’ but the reference of the designator is determined by an essential property of the referent. Thus it is not possible to say that although pain is necessarily identical with a certain physical state, a certain phenomenon can be picked out in the same way we pick out pain without being correlated with that physical state. If any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon is *pain*. (Kripke 1980, 152-153.)

Kripke selventää vielä kantaansa toisen esimerkin avulla (1980, 153–154). Luodessaan lämmön, Jumalan ei tarvinnut luoda mitään muuta kuin molekyylien liike. Mutta jotta lämpö *tuntuisi lämmöltä*, Jumalan täytyi luoda joku kokemaan se. Vasta sen jälkeen joku voi tietää, että ”lämpö on molekyylien liikettä” – kyseessä on siis a posteriori -totuus. Luodakseen C-hermosäikeiden laukeamista, Jumalan täytyi vain luoda olioita, joilla oli kyseisiä hermosäikeitä ja fysikaaliset edellytykset niiden laukeamiseen. Ei ollut merkitystä, olivatko oliot tietoisia vai eivät. Mutta jotta C-hermosäikeiden laukeamisen sai tuntumaan kivulta, Jumalan täytyi tehdä jotain ylimääräistä: antaa sen tuntua kivulta (sen sijaan että se olisi tuntunut kutinalta, lämmöltä tai ei-miltään). Ja jos tämä on todella Jumalan päätettävissä, eivät kipu ja C-hermosäikeiden laukeaminen voi olla identtisiä. Koska C-hermosäikeiden laukeamista voi esiintyä ilman kipua, ja koska ”kipu” ja ”C-hermosäikeiden laukeaminen” ovat kiinteitä nimittäjiä, tästä seuraa, että näiden kahden välinen suhde ei ole identiteetti. Siispä aivotilan ja mentaalisen tilan välinen suhde on kontingentti, ja identtisyys on ollessa suhde, joka ei voi vallita kontingenttien asioiden välillä, täytyy psykofyysisen identiteettiteorian olla väärässä.

3.2. Puuttuvan kvalian argumentti

Maailmassa jossa elämme, esiintyy tietoista kokemusta. On kuitenkin mahdollista kuvitella maailmamme kaksoiskappale – kopio, joka on täsmälleen samanlainen kuin esikuvansa kaikilta fysikaalisilta osiltaan, mutta josta puuttuu tietoinen sisältö. Mitä siis meidän maailmallamme on, mitä kopiolla ei ole? Kyseessä ei voi olla mitään fysikaalista, sillä siinä suhteessa maailmat ovat määritelmän mukaan identtisiä. Maailmallamme on siis jotain ei-fysikaalista. Siispä fysikalismi on väärässä.²⁴

Puuttuvan kvalian (*absent qualia*) yhteydessä puhutaan usein myös ihmisen kaksoisolennosta, joka on fysikaalisesti identtinen vaikkapa minun kanssani, mutta joka ei ole tietoinen – eli on ns. *filosofinen zombi*. Vaikka maailmassamme ei kenties elelekään (ainakaan kovin montaa) zombia, niin sellaisen *kuvittelemisessa* ei ole mitään ristiriitaista. Koska zombin fysikaalinen rakenne on täsmälleen sama kuin esikuvallaan, niin siitä seuraa, että tietoisuuden täytyy olla fysikaalisesta erillinen, toisin sanoen *ei-fysikaalinen* ominaisuus. Siispä fysikalismi on väärässä. Zombien *empiirisen* olemassaolon mahdollisuudella ei ole väliä – voisiko Jörn Donner olla zombi – vaan pelkkä *looginen mahdollisuus* riittää.

These systems will look identical to a normal conscious being from the third-person perspective: in particular, their brain processes will be molecule-for-molecule identical with the original, and their behavior will be indistinguishable. But things will be different from the first-person point of view. [...] And there is nothing it is like to be a zombie. (Chalmers 2002a, 249.)

Tällainen zombi on funktionaalisesti täsmälleen samanlainen kuin me.²⁵ *Psykologiset pääsytietoisuutemme* ovat samat. Ainoa mikä siltä puuttuu, on *fenomenaalinen tietoisuus*. Ruusu ei tuoksu miltään, päänsärky ei tunnu miltään, mikään ei ole *jonkinlaista*. Kvaliat puuttuvat.

²⁴ Tällaisen argumentin ovat muotoilleet alun perin Keith Campbell (1970) ja Robert Kirk (1974). Myöhemmin esim. Chalmers (1996).

²⁵ Zombi-argumentista seuraa myös muiden mielten ongelma: subjektilla ei ole mitään keinoa erottaa, ovatko kaikki muut vain zombeja.

Chalmers (1996, 97–98) perustelee zombi-kaksoisolennon *käsitteellistä mahdollisuutta* kahdella esimerkillä. Ensimmäisen esimerkin mukaan jokin suuri kansakunta, kuten vaikkapa kiinalaiset, voisi järjestäytyä siten, että se muodostaisi täsmälleen samanlaisen kausaalisen järjestelmän kuin David J. Chalmersin aivot.²⁶ Jokainen yksittäinen kiinalainen jäljittelisi yksittäisen neuronin toimintaa, ja radioyhteydet vastaisivat synapseja. Tämä voisi tapahtua robottimaisen kuoren sisällä, jota kiinalaisneuronit koordinoisivat. Ei ole merkitystä, olisiko tällainen järjestelmä tietoinen vai ei – olennaista on vain se, että on *johdonmukaista ja ristiriidatonta* ajatella, että se *ei olisi* tietoinen. Toisen esimerkin mukaan Davidilla voisi olla mikrosiruista rakennettu kaksoisolento, jonka rakenne olisi täsmälleen sama kuin esikuvallaan, paitsi siinä missä David₁ koostuisi neuroneista, koostuisi David₂ mikrosiruista. Ensimmäisen esimerkin tavoin tässäkin tapauksessa on täysin mahdollista kuvitella, että David₂ ei olisi tietoinen.²⁷

Näistä esimerkeistä voi päätyä seuraavanlaiseen argumentointiin: 1. Esimerkkien kaksoisolentojen kuvittelu ilman tietoisuutta on mahdollista. 2. Tästä johtuen myös aikaisemmin mainitun zombikaksoisolennon kuvittelu on mahdollista, sillä tietoinen kokemus ei ole loogisesti sidottu sen enempää biokemiaan kuin mikrosiruihin tai kiinalaishomunculuksiin. 3. Tietoisuus ei siis seuraa loogisesti funktionaalista rakenteesta. 4. Tietoisuus ei supervenoidu (loogisesti) fyysikaalisesta.

Argumentin rakenne vaikuttaa pätevältä, vaikkakin premissit ovat melko epätarkkoja. Onko kuviteltavuus määre, joka joko vallitsee tai ei vallitse? Entä riittääkö kuviteltavuus (looginen ristiriidattomuus), vai vaatiiko argumentti toimiakseen myös metafyyssisen mahdollisuuden? Puuttuvan kvalian argumenttia voi kritisoida myös siitä, onko fyysikaalisen kopion olettaminen ilman tietoista sisältöä ylipäätään mahdollista. Chalmers itsekin huomauttaa (1996, 99), että zombi-argumentin vastustaja voisi väittää, että jos kuvittelee fyysikaalisesti identtisen maailman, tulee automaattisesti kuvitelleeksi maailman, jossa on tietoista kokemusta.

²⁶ Esimerkki ei ole Chalmersin oma, vaan sen on esittänyt aikaisemmin mm. Ned Block (1978) (tosin kumpikaan ei, kuten arvata saattaa, puhunut juuri Chalmersin aivoista).

²⁷ Funktionalismin kannattajat olisivat tietenkin eri mieltä: heidän mukaansa sekä kiinalaishomunculusten koordinoima robotti että David₂ olisivat tietoisia, koska niillä *ex hypothesi* olisi täsmälleen samat funktiot – eli ainoa tietoisuuden vaatimus täytyisi.

Mutta siinä tapauksessa todistustaakka siirtyisi Chalmersin mukaan vastaväittäjälle, jonka pitäisi pystyä osoittamaan, missä johdonmukaiselta näyttävän argumentin ristiriitaisuus piilee. Niin kauan kuin tätä ei voida osoittaa, on pakko hyväksyä, että zombimaailma on *loogisesti mahdollinen*.

3.3. Käänteisen kvalian argumentti

Kuviteltavuus-argumentin voi muotoilla myös vain osittain: kvalioiden ei tarvitse puuttua kokonaan, vaan ne voivat olla vain erilaisia (vaikka fysikaalinen rakenne pysyykin samana). Jos kahdella oliolla on täsmälleen samanlainen fysikaalinen rakenne, mutta heidän kokemansa kvaliat ovat erilaisia, voi tästä yhtä hyvin kuin zombi-argumentistakin päätellä, ettei tietoisuus määräydy fysikaalisen rakenteen perusteella. On mahdollista kuvitella fysikaalisesti identtinen olio (vaikkapa Daniel C. Dennettin kanssa), jonka fenomenaaliset kokemukset ovat kuitenkin erilaisia kuin esikuvallaan. Ne voivat olla joko aivan uudenlaisia ennen kokemattomia kvalioita – esim. jokin kokonaan uusi väri – tai vaihtoehtoisesti käänteisiä kvalioita (*inverted qualia*) – esim. punaisen kokemus onkin vihreän kokemus ja toisin päin.

Neither would it carry any Imputation of *Falshood* to our simple *Ideas*, if by the different Structure of our Organs, it were so ordered, That *the same Object should produce in several Men's Minds different Ideas* at the same time; v.g. if the *Idea*, that a *Violet* produced in one Man's Mind by his Eyes, were the same that a *Marigold* produced in another Man's, and *vice versa*. For since this could never be known[.] (Locke 1975, 389.)

Jo 1600-luvulla Locke on esittänyt vastaavanlaisen ajatuksen käänteisen spektrin²⁸ muodossa. On mahdollista kuvitella, että Patricia näkee punaista siellä, missä Paul näkee vihreää. Heidän käytöksestään ei voi nähdä mitään eroa,²⁹ he osaavat kyllä

²⁸ Käänteinen spektri (*inverted spectrum*) on tavallisin esimerkki käänteisestä kvaliasta – argumentin voisi kuitenkin muodostaa mistä tahansa muustakin aistista. Ihmisen havaitsema värispektri on itse asiassa epäsymmetrinen, joten täydellinen käänteisyys olisi *meille* mahdotonta. Esim. Shoemaker (1982, 367) kuitenkin huomauttaa, että on täysin mahdollista *kuvitella* olio, jonka väriehavaitsemisrakenne mahdollistaisi täydellisen käänteisen värispektrin. (Tämä ei ole kuitenkaan argumentin kannalta kovin oleellista.)

²⁹ Tämä on hyvin suora behaviorismin kritiikki; tosin klassisen behaviorismin kannattajat taitavat nykyään olla vähissä. Sama toimii kuitenkin myös funktionalismia vastaan: heidän funktionaalisissa toimintoissaan ei ole mitään eroja.

toimia maailmassa juuri samalla tavalla.³⁰ Mikään neurofysiologinen fakta aivojen rakenteesta ei kerro, miksi tietty toiminto tuottaa juuri punaisen aistihavainnon (eikä vaikka sinistä tai keltaista). Punaisen havaitsemiseen liittyvät aivotilat voi siis kyllä selvittää (ja on jo hyvin pitkälle selvitettykin), mutta ne eivät kerro, miksi niihin liittyvä kvalia on juuri punaisuus. Kvalioita ei siis ole mahdollista selittää aivojen funktionaalisilla rooleilla.

I suspect that the main reason why many philosophers are hostile to the claim that such spectrum inversion is possible is that they suspect that admitting that claim will put one on a slippery slope which will eventually land one in skepticism about other minds. If I cannot know from the behavior of others that their color experiences are like my own, neither can I know that when they are cut or burned they have experiences phenomenally or qualitatively like my pains; and if I cannot know even that much, it is natural to suppose, then I cannot know anything about their minds. (Shoemaker 1982, 364.)

Vähäiseltä kuulostava värien havaitsemisen muutos johtaa siis vakaviin ongelmiin. Ei tarvitse edes kuvitella, että esimerkkihenkilö käy läpi spektrimuutoksen (kuten alan kirjallisuudessa yleensä esitetään), sillä ongelman ydin on siinä, että en edes voi tietää olenko itse havainnut syntymästani saakka värit eri tavalla kuin muut. Shoemaker päätyykin vastustamaan käänteisen kvalian argumenttia ennen kaikkea sillä perusteella, että koska skeptisismi muiden mielten suhteen on ”absurdia”, käänteisen spektrin täytyy olla mahdottomuus. Joku saattaisi vastustaa argumenttia myös jokseenkin analogisesti epifenomenalismiin päätymisen pelon takia. Tällaiset perustelut eivät ole kuitenkaan kovin vakuuttavia: sinänsä pätevältä vaikuttavia premissejä ei voi muuttaa vain johtopäätösten pelossa. Pragmaattisessa tietoteoriassa käänteinen spektri ei ole ongelma: vaikka Patricia näkee punaista siellä missä Paul vihreää, he toimivat täsmälleen samalla tavalla – joten eroa *ei ole*.

³⁰ Tästä seuraa epifenomenalismien uhka: jos he kerran toimivat täsmälleen samalla tavalla, onko kvalialla vaikutusta mihinkään? Kenties hyvin syvällisellä haastattelulla eroja kuitenkin olisi mahdollista selvittää, sillä kuten jo mainitsin, ihmisen havaitsema värispektri ei ole symmetrinen (jossain värissä on mahdollista havaita enemmän sävyjä kuin jossain toisessa).

Eräät fysikalismien kannattajatkin ovat myöntäneet, että on mahdollista, että henkilön värikokemukset voisivat olla käänteisiä, vaikka fysikaalinen rakenne olisi sama.³¹ Chalmers huomauttaakin (1996, 101), että käänteisen kvalian argumentti on heikompi kuin puuttuvan kvalian argumentti, sillä on johdonmukaista ajatella, että käänteinen spektri on mahdollinen, mutta zombit eivät. Silloin tietoisuus olisi redusoitavissa fysikaaliseen, vain jotkut sen ominaisuudet eivät olisi. Molemmista argumenteista seuraa kuitenkin sama lopputulos: sekä zombimaailma että käänteinen spektri ovat käsitteellisesti koherentteja, loogisesti mahdollisia vaihtoehtoja.

3.4. Seuraako kuviteltavuudesta metafyyminen mahdollisuus?

Ontologinen argumentti fysikalismia vastaan vaatii, että kuviteltavuudesta seuraa metafyyminen mahdollisuus. Episteemiselle argumentille sen sijaan riittää looginen mahdollisuus.³² Vaikka hyväksyttäisiinkin, että puuttuvan tai käänteisen kvalian kuviteltavuudesta seuraa looginen mahdollisuus, argumentista nousee siis uusi kysymys: seuraako loogisesta mahdollisuudesta edelleen metafyyminen mahdollisuus? Tällöin loogisella mahdollisuudella ei olisi enää juurikaan merkitystä, joten kysymyksen voi yksinkertaistaa: seuraako kuviteltavuudesta metafyyminen mahdollisuus?³³ Asiasta ei vallitse yksimielisyyttä. Vastustajat ovat usein vedonneet kriipkeläiseen välttämättömän a posteriorin käsitteeseen.

We have concluded that an identity statement between names, when true at all, is necessarily true, even though one may not know it *a priori* (Kripke 1980, 108).

Lause ”Vesi on H₂O:ta” on metafyymsisesti välttämättä tosi, mutta se ei ole a priori - totuus, sillä olisi mahdollista, että vesi olisi osoittautunut olevan jotain muuta,

³¹ Martine Nida-Rümelin (1996) on tuonut esiin eräänlaisen tosielämän version käänteisestä kvaliasta. Tutkimukset (miesten) värinäöstä ovat osoittaneet, että jokaista 10 000 miestä kohti 14 näkee punaista siellä, missä muut vihreää ja toisin päin. Heidän aivotoinnoissaan on tosin havaittavissa eroavaisuus, joten huomio ei sinänsä tue käänteisen kvalian argumenttia, vaan on päinvastoin täysin yhteensopiva fysikalistisen funktionalismin kanssa.

³² Tarkoitan loogisella mahdollisuudella *käsitteellistä* tai *episteemistä* mahdollisuutta. X on loogisesti mahdollinen silloin, kun se ei ole loogisesti ristiriitainen. Toisin sanoen, kun X on kuviteltavissa, niin se on loogisesti mahdollinen.

³³ On syytä muistaa, että kyseessä on juuri *metafyyminen*, ei fyysinen mahdollisuus. Kukaan ei ole varmaankaan väittänyt, että kuviteltavuudesta seuraisi fyysinen mahdollisuus. On esimerkiksi kuviteltavissa, että Patricia voisi matkustaa Andromedan galaksille. Jos siitä seuraa mahdollisuus, niin se on metafyyminen – ei fyysinen mahdollisuus.

vaikkapa XYZ:aa. On siis *loogisesti* mahdollista, että vesi ei olisi H₂O:ta, vaikka se ei olekaan *metafyysisesti* mahdollista. Koska vesi kuitenkin aktuaalisessa maailmassa on H₂O:ta, niin Kripken teesin mukaan vesi on H₂O:ta kaikissa metafyysisesti mahdollisissa maailmoissa (koska molemmat ovat kiinteitä nimittäjiä). Siispä lause ”Vesi on H₂O:ta” on *välttämättä* tosi, vaikkakin se on tiedettävissä vain *a posteriori*. Tilanne voisi olla sama zombien kohdalla: kenties ne ovat mahdollisia loogisesti, mutta eivät metafyysisesti. Chalmersin ontologinen fysikalismen kritiikki siis kaatuisi, ja kuilu tietoisuuden ja fysikaalisen välissä olisi korkeintaan episteeminen.

Chalmersilla on kuitenkin vastaus (tähänkin) vastaväitteeseen.³⁴ Selvittääkseen, seuraako kuviteltavuudesta mahdollisuus, on ensin selvennettävä kuviteltavuuden (jokseenkin häilyvää) käsitettä. Chalmersin mukaan kuviteltavuudesta (oikein muodostettuna) seuraava mahdollisuus on metafyysinen. Chalmers (2002b, 145–159) muotoilee kolme kuviteltavuuden käsitteen ulottuvuusparia, joiden yhteinen kokoonpano määrää, kuinka vahva silta kuviteltavuudesta mahdollisuuteen muodostuu.

a) Prima facie vs. ideaalinen kuviteltavuus

Lause L on prima facie -kuviteltavissa silloin, kun se vaikuttaa ensisilmäyksellä (päällisin puolin tarkasteltuna) kuviteltavalta. L on ideaalisesti kuviteltavissa silloin, kun se läpäisee tarkemmankin tarkastelun: se on kuviteltavissa pätevien rationaalisten syiden takia. Prima facie -lauseen mahdollisuus voi tarkemmassa tarkastelussa kumoutua tai saada vahvistusta ja muuttua ideaaliseksi kuviteltavuudeksi, joka on siis ensin mainittua vahvempi oikeutus mahdollisuudelle.

b) Positiivinen vs. negatiivinen kuviteltavuus

Positiivisessa kuviteltavuudessa muodostetaan *mielikuvituksen* avulla jokin käsitys tilanteesta, jossa pätee L: millä tavalla objektit ja ominaisuudet olisivat järjestyneet kyseisessä tilanteessa. Jos tämä (mieli)kuvittelu on mahdollista, L:n

³⁴ Vastausta ei voi kutsua kovin yksinkertaiseksi. Syvälinen paneutuminen siihen vaatisi enemmän kuin kokonaisen pro gradu -työn. Otan sen tässä kuitenkin lyhyesti esiin, koska se liittyy olennaisesti kuviteltavuuden ja mahdollisuuden suhteeseen.

mahdollisuus on verifioitu. Kuviteltu tilanne siis määrää, toteutuuko positiivinen kuviteltavuus. Yksisarvisen kuvittelemisen kaikissa yksityiskohdissaan ei takaa kuvitellun yksisarvisen olemassaoloa, mutta se takaa sen käsitteellisen kuviteltavuuden intentionaalisen objektina. Jos L kuitenkin sisältää apriorisen ristiriidan, ei se ole lainkaan kuviteltavissa: esimerkiksi ristiriitaisin ominaisuuksin varustettua geometristä kuviota ei voi kuvitella (ellei sitä ole kuviteltu virheellisesti ilman kaikkia yksityiskohtia). Negatiivisessa kuviteltavuudessa L taas on kuviteltavissa, jos L:ää ei voi sulkea pois (apriorisesti, tai siten, ettei L:lle ole selkeää ristiriitaa). Jos L on positiivisesti kuviteltavissa, se on myös negatiivisesti kuviteltavissa, koska jos voi johdonmukaisesti kuvitella tilanteen, joka verifioi L:n, niin L:ää ei voi sulkea pois. (Sama ei aina toimi toisin päin.) Positiivinen kuviteltavuus siis on vahvempi argumentti mahdollisuuden puolesta kuin negatiivinen kuviteltavuus.

c) Primaarinen vs. sekundaarinen kuviteltavuus

Kolmas erottelu perustuu Chalmersin kaksiulotteiseen semantiikkaan (*two-dimensional semantics*), joka liittyy kripkeläiseen a posteriori -välttämättömyyteen. Primaarinen kuviteltavuus on episteeminen kanta, joka perustuu a priori -päätelmiin. Tällä tavalla ”Vesi ei ole H₂O:ta” on kuviteltavissa, koska väite ei sisällä apriorista totuutta, vaan vesi voisi olla myös XYZ:aa. Chalmers (2002b, 157) kutsuu tällaista *episteemiseksi mahdollisuudeksi*. Lause L on primaarisesti (eli *episteemisesti*) kuviteltavissa, kun on kuviteltavissa että L todella, *aktuaalisesti*, vallitsee. Kysymys kuuluu siis: mitä jos aktuaalinen maailma on jonkinlainen? Koska primaarinen kuviteltavuus on kuitenkin aina apriorista, ei aktuaalisen maailman empiirisellä tilanteella ole väliä.

L on taas sekundaarisesti (eli *subjunktiivisesti*) kuviteltavissa, kun *olisi voinut olla* niin, että L. Voimme kuvitella monenlaisia kontrafaktuaalisia asiointiloja, jotka olisivat voineet vallita, mutta eivät vallitse. Sekundaarinen kuviteltavuus on aposteriorista. Tällä tavalla ei siis ole kuviteltavissa, että ”Vesi ei ole H₂O:ta”, sillä on tarkastettava, miten asiat todellisessa maailmassa ovat; ja empiirisen tutkimuksen jälkeen väite ei ole enää kuviteltavissa. Ei siis ole koherentisti kuviteltua kontrafaktuaalista tilannetta, joka vahvistaisi, että ”Vesi ei ole H₂O:ta”. Vaikka

kuvittelisimme, että vesi on XYZ:aa, niin empiirinen tarkastelu paljastaa *a posteriori*, että XYZ ei ole vettä, vaan korkeintaan *vesimäistä ainetta*. Primaarinen kuviteltavuus on sekundaarista vahvempi edellytys mahdollisuudelle (modaaliselle mahdollisuudelle, josta tässä on kyse), sillä vaikka sekundaarinen kuviteltavuus johtaisikin mahdollisuuteen, se ei voi johtaa a priori -modaali-teetteihin aposteriorisen luonteensa takia.

Kuviteltavuus muodostuu siis käytännössä näiden tapojen yhdistelmästä: prima facie negatiivinen sekundaarinen kuviteltavuus ei ole kovin vahva edellytys mahdollisuudelle, mutta ideaalinen negatiivinen primaarinen on jo vahvempi yritys. Vahvimman sillan kuviteltavuudesta mahdollisuuteen muodostaa Chalmersin mukaan (2002b, 171) kuitenkin *ideaalinen positiivinen primaarinen kuviteltavuus*. Jos L on kuviteltavissa tällä tavalla, se johtaa Chalmersin mukaan ”lähes täysin varmasti” *primaariseen* mahdollisuuteen. Sekundaarinen kuviteltavuus johtaa sekundaariseen mahdollisuuteen (eli empiiriseen mahdollisuuteen), ja primaarinen kuviteltavuus primaariseen mahdollisuuteen (metafyysiseen mahdollisuuteen). Kuviteltavuus, kunhan se on oikein muodostettu, johtaa siis aprioriseen metafyysiseen mahdollisuuteen. Jos Chalmers on oikeassa, niin modaaliargumentti osoittaa fysikalismiin (episteemisen lisäksi) myös ontologisesti vääräksi.³⁵

Chalmers ei väitä, että mikä tahansa, mitä voimme kuvitella oikealla (ideaalisella positiivisella primaarisella) tavalla, voisi olla metafyysisesti mahdollista. Lause ”Vesi ei ole H₂O:ta” on primaarisella tavalla kuviteltavissa, joten se on vain episteemisesti mahdollinen. Sekundaarisesti se ei ole kuviteltavissa, koska tällä tavalla kuviteltu tilanne (esim. ”Vesi on XYZ:aa”) on kontrafaktuaali, joka ei toteudu aktuaalisessa maailmassa, jonka luonne on aposteriorisen tarkastelun jälkeen saatu selville. Jos zombien kuviteltavuus osoittautuu analogiseksi (*kuviteltavissa vain primaarisesti, mutta ei sekundaarisesti*), niin Chalmersin ontologinen argumentti fysikalismia vastaan kaatuu. Mutta tietoisuus (eli sen puutteen kuvittelemisen, toisin sanoen zombien kuvittelemisen) ei Chalmersin (2002a, 256) mukaan kuitenkaan

³⁵ Chalmers huomauttaa (2002b, 172), että väitelauseiden kuviteltavuus sekä mahdollisuus ovat lisäksi riippuvaisia siitä, kuka lauseet esittää. Jos lause L on *kuviteltavissa* jollekin puhujalle, L on *mahdollinen* tuolle puhujalle. Lause ”Minä en ole Johanna Hyvärinen” on aivan eri tavalla *kuviteltavissa* kirjoittajalle kuin lukijalle, samoin se on aivan eri tavalla *mahdollinen* kirjoittajalle ja lukijalle. Tällainen puhuja-relatiivisuus ei kuitenkaan vaikuta kuviteltavuus–mahdollisuus -suhteeseen, koska se koskee kumpaakin tasapuolisesti.

toimi vastaavalla tavalla, sillä sen sekundaarista kuviteltavuutta ei voi rajata pois samoin kuin veden tapauksessa. Chalmers käyttää *kaksiulotteista semantiikkaa* apuvälineenä argumentoidessaan, miksi tietoisuuden suhteen kuviteltavuudesta seuraa mahdollisuus muista maailman ilmiöistä poiketen.

Kaksiulotteisessa semantiikassa käsitteillä on kaksi intensiota. *Primaarisessa intensiossa* referenssi määräytyy aktuaalisen maailman perusteella riippuen siitä, millaiseksi maailma osoittautuu. Jos se osoittautuu yhdenlaiseksi, käsite poimii jonkin asian, jos se osoittautuu toisenlaiseksi, käsite poimii jonkin toisen asian. *Sekundaarisessa intensiossa* referenssi määräytyy kontrafaktuaalisten maailmojen perusteella. Tässä tapauksessa aktuaalisen maailman luonne on jo lyöty lukkoon, ja kysymys kuuluu, *mikä olisi ollut* tilanne. (Chalmers 1996, 57.) Jos esim. Patricia kuvittelee, että vesi ei ole H₂O:ta, niin hän ei tarkalleen ottaen kuvittelekaan ”vettä” vaan vain jotain ”vesimäistä ainetta”. Voi siis sanoa, että joko hän kuvailee kuvittelun kohteen väärin, tai sitten hän kuvailee kohteen *primaarisen intension*. Mutta kun hän puhuu ”vedestä”, hän ei voi tarkoittaa sillä mitään muuta kuin ”H₂O:ta” (*sekundaarinen intensio*).

[W]e might say that the primary intension picks out the dominant clear, drinkable liquid in the oceans and lakes; or more briefly, that it picks out the *watery stuff* in a world. However, given that “water” turns out to refer to H₂O in the actual world, Kripke notes (as does Putnam) that it is reasonable to say that water is H₂O in every counterfactual world. The *secondary intension* of “water” picks out the water in every counterfactual world; so if Kripke and Putnam are correct, the secondary intension picks out H₂O in all worlds. (Chalmers 1996, 57.)

Jos aktuaalisessa maailmassa W lause S on totta, niin W *verifioi* S:n. Jos kontrafaktuaalisessa maailmassa W lause S on totta, niin W *täyttää ehdon* (*satisfies*) S. Verifikaatio vaatii siis maailmojen episteemistä tarkastelua, ja ehdon täyttäminen maailmojen kontrafaktuaalista tarkastelua. Esimerkiksi jos S on ”Vesi ei ole H₂O:ta”, ja W on XYZ-maailma, niin W verifioi S:n, mutta W ei täytä ehtoa S. Toisin sanoen primaarisesti S on totta W:ssä, mutta sekundaarisesti epätotta. Kun S on kuviteltavissa, on siis ainakin jokin maailma W, joka verifioi sen. (Muuten S ei

olisi kuviteltavissa.) Tällöin *S* on siis *episteeminen mahdollisuus*. Jos tämä *W* osoittautuu aktuaaliseksi maailmaksi, niin selviää, että *S*. (Chalmers 20002a, 256).

Tietoisuuden suhteen kuviteltavuudessa on kysymys fysikaalisen kuvittelemisesta ilman fenomenaalista. Vaikka yleensä (kuten veden tapauksessa) verifioitumisesta (episteemisestä mahdollisuudesta) ei voi johtaa suoraan ehdon täyttymistä, niin tietoisuus muodostaa poikkeuksen säännöstä – tietoisuuden käsitteen luonteesta johtuen. Chalmers viittaa Kripken ajatukseen siitä, että tietoisuuden käsitettä ei voi jakaa samalla tavoin kuin veden käsitteestä voi erottaa vesimäisen aineen. Kipua ei voi erottaa kivun tuntemuksesta. ”All it means to be a conscious experience, in any possible world, is to have a certain feel” (Chalmers 1996, 133). Jos *W* verifioi, että ”on tietoisuutta”, niin *W* täyttää ehdon, että ”on tietoisuutta” ja päinvastoin. Sillä jos *W* verifioi, että ”on tietoisuutta”, niin se sisältää jotain, joka ainakin *tuntuu* tietoisuudelta. Ja jos jokin *tuntuu* tietoisuudelta, se *on* tietoisuutta *kontrafaktuaalisessakin maailmassa*. (Chalmers 2002a, 256.)

Chalmersin mukaan siis puuttuvan ja käänteisen kvalian kuviteltavuudesta seuraa, että ne ovat todellisia metafysisiä mahdollisuuksia. Ja *mahdollisuus* todistaa, että argumentit osoittavat ontologisen fysikalismien olevan väärässä, eli todellisuuteen kuuluu ei-fysikaalisia ominaisuuksia.³⁶

Chalmersin päättely kuviteltavuudesta metafysiseseen mahdollisuuteen on mielenkiintoinen argumentti. Mutta onko se pätevä? Se perustuu vahvasti tietoisuuden luonteen erityislaatuisuuteen, sekä siihen, että tietoisuus ei ole periaatteessakaan selitettävissä samalla tavalla kuin muut maailman ilmiöt. Mutta seuraako selityksen puutteesta todella ontologinen fakta, kuten Chalmers väittää? Omasta mielestäni selityksen puute ei vielä itsessään oikeuta muihin johtopäätöksiin kuin – selityksen puutteeseen. Voidaan myös kysyä, onko Chalmersin koko jaottelu primaariseen ja sekundaariseen intensioon oikeutettu. Hän huomauttaa itsekin:

[O]ne can argue that this view leads to an underlying *modal dualism*, with independent primitive domains of logical and

³⁶ Kysymys kuviteltavuuden ja metafysisen mahdollisuuden suhteesta on hyvin mielenkiintoinen, ja vaatisi paljon laajempaa tutkimusta, kuin mikä on tämän työn puitteissa mahdollista.

metaphysical possibility; and one can argue that this is unacceptable (Chalmers 2002a, 257).

3.5. Moninaisen toteutuvuuden argumentti

Kolmas kuviteltavuudesta lähtevä argumentti on moninaisen toteutuvuuden (*multiple realizability*) argumentti, joka ei ole suunnattu funktionalismia vastaan (*puuttuvan ja käänteisen kvalian* argumenttien tavoin), vaan ennen kaikkea tyyppi-identiteettiteoriaa sekä yleisemmin (fysikalistista) reduktionismia vastaan. On mahdollista kuvitella, että neuraalinen tila, joka aiheuttaa kvalian Paulissa, on eri kuin neuraalinen tila, joka aiheuttaa saman kvalian Patriciassa tai pelikaanissa. Argumentin ovat esittäneet alun perin Hilary Putnam ja Jerry Fodor 1960- ja 1970-luvuilla³⁷ (Funkhouser 2007, 303). Heidän pyrkimyksensä oli kumota identiteettiteoria ja tuoda esiin omaa kantaansa – funktionalismia.

Tyyppi-identiteettiteorian mukaan mentaalinen tila, esimerkiksi kipu, on identtinen fysikaalisen aivotilan, esimerkiksi C-hermosäikeiden laukeamisen kanssa. Moninaisen toteutuvuuden argumentin mukaan sama mentaalinen tila voi kuitenkin toteutua muidenkin neuraalisten syiden seurauksena. Kipu *toteutuu moninaisesti* (tai *monitoteutuu*), jos on *eri tapoja*, joilla oliot voivat kokea *samanlaista* kipua. Putnam ja Fodor väittivät, että näin todella tapahtuu: heidän mukaansa on useita tapoja, joilla kipu ja muut mentaaliset tilat voivat toteutua fysikaalisesti. Mentaaliset tilat eivät siis voi olla identtisiä fysikaalisten tilojen kanssa. Esimerkiksi Putnam (1967) huomautti, että vaikka ihmisten kokema kipu redusoituisikin C-hermosäikeiden laukeamiseen, näin ei kuitenkaan välttämättä tapahdu muilla eläimillä. Nisäkkäiden, matelijoiden ja lintujen (tai marsilaisten) fysikaalinen rakenne ei ole sama – vaikka voi kuitenkin olettaa, että ne kaikki tuntevat kipua. Siispä identiteettiteoria on väärässä, ja mentaalinen ei redusoidu fysikaaliseen.³⁸

³⁷ Putnam (1967) ja Fodor (1974).

³⁸ Nykyisin moninaisen toteutuvuuden esimerkkinä käytetään usein *propositionaalisia asenteita* (uskomuksia, haluja jne.) eli toisin sanoen puhutaan psykologiasta, eikä enää *mentalisesta* tai varsinkaan kvalioista (Bickle 2008). Tässä työssä kipu toimii kuitenkin parempana esimerkkinä, sillä se on fenomenaalinen kokemus *par excellence*.

Argumenttia voi vastustaa kahdelta suunnalta: joko kieltää sen premissi (*mentaalinen toteutuu moninaisesti*) tai siitä seuraava johtopäätös (*mentaalinen ei redusoidu fyysikaaliseen*). Funkhouser (2007) esittelee kaksi tapaa, joilla premissiä on kritisoitu, ja Jaegwon Kim taas kritisoi kahdella tavalla johtopäätöstä; itse puolestani kritisoin näitä kritiikkejä.

Premissin kritiikki 1: On mahdollista väittää, ettei kyseessä todellisuudessa ollutkaan sama kipu, vaikka siltä saattoikin vaikuttaa.³⁹ Jos väite pitää paikkansa (Patrician ja mustekalan kipu ei olekaan samanlaista), ei kipujen fyysikaalisella moninaisella toteutumisella ole väliä, ja identiteettiteoria pelastuu. Selvittääkseen toimiiko kritiikki, olisi selvitettävä, mitä on samanlainen kipu, tässä tapauksessa ihmisten ja mustekalojen välillä. (Behavioristeilla voisi olla tähän keinonsa, mutta behaviorismi ei vastaa tämän työn kannalta relevantteihin kysymyksiin tietoisuuden suhteen.) Funkhouser ehdottaa käsitteellistä ja *empiiristä* tutkimusta:

Purely conceptual considerations can give us a proto-theory to start with [...] empirical work is required to see if humans and octopi can satisfy the same psychological kind, though differing in their realizations. (Funkhouser 2007, 305.)⁴⁰

Mutta mikä olisi sellainen empiirinen keino, jolla voisi selvittää, onko ihmisten kipu ja mustekalojen kipu samanlaista? Fyysikaalista toteutusta voi tutkia empiirisesti vaikka molekyyalitasolle saakka, mutta siitähän tässä juuri ei ole kyse. Tässä on syytä muistaa Nagelin huomio siitä, ettemme voi koskaan tietää, miltä tuntuu olla jokin muu kuin itse on. Emme saa koskaan tietää, millaista mustekalojen kokema kipu on, joten Funkhouserin ratkaisuehdotusta ei ole mahdollista käytännössä toteuttaa. Väitän siis, että ensimmäinen kritiikki on mieletön, sillä kivuntuntemuksen väitettyä erilaisuutta ei voi yksinkertaisesti tietää.

Tässä on tosin syytä huomata, että sama kritiikki koskee tosin myös itse moninaisen toteutuvuuden argumenttia: argumentin vaatimaa kivun samanlaisuutta ei voi myöskään tietää (samoista syistä kuin erilaisuuttakaan). Moninaisen toteutuvuuden

³⁹ Tällaisen kritiikin on esittänyt mm. Thomas Polger (Funkhouser 2007).

⁴⁰ Funkhouser puhuu ”psykologisesta puolesta” (*psychological kind*), ei edes mentaalista, saati kokemuksen fenomenalisuudesta, kvalitatiivisuudesta tai vain yksinkertaisesti kvaliasta. Tämä on jokseenkin ongelmallista, koska hän käsittelee kuitenkin Putnamin ja Fodorin kipu-esimerkkiä, mutta ei siis tule ottaneeksi huomioon subjektiivisen kokemuksen erityislaatuisuutta.

argumenttikin suistuu siis muiden mielen ongelmaan. Jos haluaa kuitenkin *uskoa* (terveen järjen mukaisesti), että muita mieliä on, ja että myös eläimet tuntevat kipua, voittaa moninaisen toteutuvuuden argumentti ensimmäisen kritiikkinsä.⁴¹

Premissin kritiikki 2: Jos hyväksytään, että kipu on sama, voidaan kuitenkin puolustautua väittämällä, että myös fyysikaalinen toteutus on (ainakin riittävän) sama. Ovatko ihmisen ja mustekalan hermosäikeet, solut ja aivot siis riittävän samanlaisia kiputilanteiden ollessa kyseessä? Tämä riippuu tietenkin erottelutavan tarkkuudesta. William Bechtel ja Jennifer Mundale ovat väittäneet (Funkhouser 2007, 305), että moninaisen toteutuvuuden argumentti vaikuttaa uskottavalta vain siksi, että käytämme erilaisia tapoja erottaessamme mentaalisia ja neurologisia puolia. Oli erottelun tarkkuus millainen tahansa (hienojakoinen tai karkeajakoinen), on tarkkuuden oltava sama sekä mentaalisella että neurotasolla, jotta mitkään johtopäätökset olisivat oikeutettuja. Jos fyysikaalista toteutusta tulkitaan karkealla tavalla (ihmisen ja mustekalan neurologia kiputilanteissa on käytännössä samanlaista), moninaisen toteutuvuuden argumentti kumoutuu. Jollain toisella tarkkuuden asteella luokiteltuna niiden neurologia on toki erilainen – mutta toisenlaisella tarkkuudella niiden mentaalinen kipukin on erilaista.

Kritiikki on mielestäni onnistunut siinä mielessä, että se osoittaa erottelun tarkkuusasteen merkityksen argumentin toimivuudelle. Tavoitteessaan – moninaisen toteutuvuuden argumentin kumoamisessa ja tyyppi-identiteettiteorian pelastamisessa – se ei kuitenkaan onnistu. Ihmisen aivot ja mustekalan aivot *eivät ole identtiset* tyyppi-identiteettiteorian vaatimassa mielessä, sillä tuntuisi omituiselta väittää, että kipu *on identtinen* ihmis-/mustekala-/simpukka-/kirahvi- jne. aivosysteemin kanssa. Lajien fyysikaalinen aivorakenne ei ole yksi *tyyppi*, vaan *useita tyyppejä*, ja kvalia monitoteutuu. Sen sijaan token-identiteettiteorian tämä kritiikki voi pelastaa. Ihmisen kipu redusoituu ihmisen aivoihin, mustekalan kipu mustekalan aivoihin. Mutta kipu itsessään (*tyyppi*) ei redusoidu.⁴²

⁴¹ Behaviorismi – vaikka kykenemätön selittämään monia tietoisuuteen liittyviä kysymyksiä – on kuitenkin jossain tarkoituksessa hyödyllinen oppi: muiden käytöksestä on uskottavaa tehdä päätelmiä näiden tietoisista kokemuksista.

⁴² Sittemmin on esitetty, että realisaatio voi vaihdella myös lajin sisällä ja jopa yksilön sisällä, esim. tilanteessa, jossa henkilön aivot ottavat plastisuutensa ansiosta haltuunsa uusia tehtäviä vaikkapa onnettomuuden tai leikkauksen jälkeen. Tällaista kantaa on kutsuttu *radikaaliksi* moninaiseksi toteutuvuudeksi. (Bickle 2008).

Johtopäätöksen kritiikki 1: Vaikka hyväksyisikin moninaisen toteutuvuuden argumentin premissin, eli mentaalisten ilmiöiden fysikaalisen monitoteutumisen, voi silti väittää, että siitä ei seuraa reduktion mahdottomuus – mentaalinen voi silti redusoitua fysikaaliseen. Jaegwon Kim (1992) on esittänyt, että monien erilaisten fysikaalisten systeemien muodostama disjunktio (esim. ihmisneurologia V mustekalaneurologia V simpukkaneurologia V marsilaisneurologia jne.) on todella olemassa oleva tyyppi tai laji, joten mentaalinen redusoituu yhteen, eikä moneen lajiin. Kim (1992, 140) selittää tarkoitustaan *jade*-esimerkin avulla. Aikaisemmin luultiin, että jade on oma mineraalilajinsa. Voitiin muodostaa laki (L) Jade on vihreää. Myöhemmin on kuitenkin selvinnyt, että todellisuudessa jade voi olla rakenteeltaan kahta keskenään täysin eri mineraalia: jadeiittia tai nefriittia. Laki (L) koostuukin siis lauseista (L₁) Jadeiitti on vihreää ja (L₂) Nefriitti on vihreää. Kimin mukaan (L) voi silti olla itsenäinen laki: jos jokin on jadea (siis jadeiittia tai nefriittia), niin se on vihreää. Jade on siis esimerkki todesta disjunkttiivisesta lajista (*true disjunctive kind*). Samalla tavalla kivun anatomia, vaikka koostuukin monien lajien kipujen anatomioista, voi olla oma lajinsa.

Kimin esittämään disjunktio-periaatteeseen on helppo vastata: tällaiset disjunktion avulla muodostetut ”lajit” eivät ole mitään todellisia lajeja. Ne ovat liian epäyhtenäisiä: ollakseen laji, pitäisi olla tarpeeksi samanlaisuutta tai yhtenäisyyttä, mutta ainoa mikä näitä yhdistää, on mentaalinen tila, jonka ne tuottavat. Ne toimivat siis juuri moninaisen toteutuvuuden argumentin kuvailemalla tavalla, joten disjunktio-periaate on *ad hoc* -muotoinen päättely.

Johtopäätöksen kritiikki 2: Kimin toinen ratkaisu on vedota *lokaaliseen reduktioon* (johon myös premissin toinen kritiikki päättyi): ihmiskipu redusoituu ihmisneurologiaan ja simpukkakipu simpukkaneurologiaan. Kim väittää, että tällä tavalla on mahdollista säilyttää reduktiivinen identiteettiteoria. Lokaalisen reduktion ongelma on kuitenkin siinä, että vaikka simpukkakipu redusoituisikin simpukkaneurologiaan, ei itse kipu *lajina* tai *tyyppinä (kind)* redusoidu. Kim kuitenkin ratkaisee ongelman kieltämällä koko lajien olemassaolon.

[P]ain as a property or kind must go. Local reduction after all is reduction, and to be reduced is to be eliminated as an *independent* entity. You might say: global reduction is different in that it is also *conservative* – if pain is globally reduced to physical property P, pain survives as P. But it is also true that under local reduction, pain survives as N_h in humans, as N_r in reptiles, and so on. It must be admitted, however, that pain as a kind does not survive multiple local reduction. But is this so bad? (Kim 1992, 146.)

Kivun käsitteen ei Kimin mukaan tarvitse viitata mihinkään todelliseen olemassa olevaan lajiin – samalla tavalla kuin jadekaan ei viittaa maailmassa olevaan todelliseen lajiin, vaikka jaden käsite on olemassa. Kim (1992, 147) erottaakin siis toisistaan *käsitteet* ja *ominaisuudet* (eli lajit). Voimme käyttää kivun käsitettä, mutta se ei viittaa mihinkään objektiiviseen lajiin sen enempää kuin jadekaan. Koska *kipua itsessään* ei ole, ei se tietenkään myöskään redusoidu mihinkään, ja moninaisen toteutumisen argumentti on kumottu.

But is this second approach a form of mental eliminativism? In a sense it is: as I said, on this approach no properties in the world answer to general, species-unrestricted mental concepts. (Kim 1992, 147.)

On olemassa ihmiskipua, simpukkakipua, marsilaiskipua – mutta ei kipua *simpliciter*. Kim ei siis väitä joidenkin eliminativistien lailla, ettei mentaalisia tiloja olisi lainkaan olemassa. Ne ovat kyllä olemassa, mutta samalla tavalla kuin jade, jonka käsite ei viittaa todelliseen ominaisuuteen. Kukaan ei kuitenkaan kiellä, etteikö jadea olisi olemassa.

Tähän kritiikkiin vastatakseen täytyisi tutkia hyvin syvällisesti käsitteiden ja ominaisuuksien metafysiikkaa. Eliminativismin vastustaja voisi sanoa, että *lajit* – kuten kipu itsessään – ovat kuitenkin tärkeä osa teorianmuodostusta, ja siinä mielessä olemassa todellisuudessa. Kim huomauttaa (1992, 148), että jos haluaa säilyttää globaalin psykologian, joka pätee kaikille lajeille ja järjestelmille, voi aina päätyä kannattamaan hänen esittämänsä disjunktio-mallia, joka säilyttää mentaaliset ominaisuudet ominaisuuksina itsessään. Oli lajien asema mikä tahansa, reduktiota Kim ei kuitenkaan kykene pelastamaan, sillä lokaalinen reduktio ei ole tyyppi-identiteettiteorian vaatimusten mukaista reduktiota, jos reduktiota lainkaan.

Voidaan siis todeta, että moninaisen toteutuvuuden argumentin kritiikit eivät siis onnistu kumoamaan sitä, vaan on todettava, että mentaaliset ominaisuudet voivat toteutua monin eri fysikaalisin tavoin. Psykofyysinen identiteettiteoria on siis väärässä. Fysikalismia argumentti ei kuitenkaan kumoa – reduktiivisen fysikalismin kylläkin.

3.6. Modaaliargumentin tulokset

Koko modaaliargumentti on hyvin tehokas argumentti fysikalismia vastaan, sillä se hyökkää sitä vastaan joka suunnalta. Jos kannattaa identiteettiteoriaa, niin Kripke ja moninaisen toteutuvuuden argumentti pakottavat kääntymään funktionalismin puoleen. Kun on funktionalisti, puuttuvan- ja käänteisen kvalian argumentit pakottavat kääntymään takaisin identiteettiteoriaan. Kehästä pääsee ulos kieltämällä kvaliat kokonaan, mutta jos on valmis siihen, ei tietoisuuden kanssa alun perinkään ollut mitään ongelmaa.

Kuviteltavuuteen (ja siitä johdettuun mahdollisuuteen) perustuvat argumentit osoittavat, että mentaalinen ja fysikaalinen eivät ole loogisesti sidoksissa toisiinsa. Puuttuvan kvalian ja käänteisen kvalian argumentit osoittavat, että fenomenaliset ja fysikaaliset ilmiöt ovat kuviteltavissa toisistaan erillisinä. Koko argumentti perustuu juuri *kuviteltavuuteen*, joka tosin on jokseenkin monimerkityksinen käsite. Kuviteltavuudesta seuraa looginen mahdollisuus, sillä kuviteltavuus *vaatii toteutuakseen* loogisen mahdollisuuden. (Ilman loogista mahdollisuutta olevaa asiaa on mahdotonta kuvitellaakaan.) Zombien kuvittelemineen on siis *loogisesti mahdollista*. Tämä ei kuitenkaan vielä kerro, onko se myös metafyyysisesti mahdollista. Jo loogisesta kuviteltavuudesta seuraa kuitenkin episteeminen kuilu fenomenalisten ja fysikaalisten ominaisuuksien välillä, sillä se osoittaa, että mikään fysikaalisessa rakenteessa ei selitä, miksi kokemukset tuntuvat juuri jonkinlaisilta.

Jos kuviteltavuudesta seuraa metafyyysinen mahdollisuus (kuten Chalmers väittää), niin supervenienssiteoria on väärässä, sillä on mahdollista olla mentaalinen ero ilman fysikaalista eroa. Tällöin todellisuuteen kuuluu siis jotain ei-fysikaalista. Kysymys

kuuluukin: seuraako kuviteltavuudesta metafyyssinen mahdollisuus? Chalmersin argumentoinnin perusta on oletus fenomenaalisten ominaisuuksien erityislaatuudesta. Tämän oletuksen pohjalta hän johtaa ontologisen eron, vaikka mielestäni erityislaatuisuus ei vielä itsessään osoita muuta kuin episteemisen eron. Chalmersin väite (sekä ylipäättään kuviteltavuuden ja mahdollisuuden suhde) vaatisi kuitenkin lisätutkimusta.

Kolmas kuviteltavuudesta lähtevä argumentti, moninaisen toteutuvuuden argumentti, on hyvin tehokas tapa osoittaa, että reduktiivinen fysikalismi (kuten identiteettiteoria) on väärässä. Ajatus, että mentaalilla tiloilla voi olla monenlaisia fysikaalisia toteutuksia, on hyvin intuitiivinen ja luonteva. Reduktionistien vastaukset tuntuvatkin keinotekoisilta. Tämän argumentin tapauksessa myös empiirinen todistusaineisto osoittaa, että mentaalinen/psykologinen tila voi toteutua moninaisesti: aivojen plastisuus mahdollistaa uudenlaisten roolien syntymisen.⁴³ Siispä mentaalinen tila ei voi olla identtinen tietyn aivotilan kanssa, koska se voi toteutua muidenkin aivotilojen kanssa.

Modaaliargumentista saattaa herätä kysymys, millä tavalla olisi mahdollista saada tietää, onko jollain oliolla fenomenaalisia kokemuksia (vai onko olio zombi). Koska fenomenaalisiin kokemuksiin ei ole mahdollista päästä käsiksi kuin ensimmäisen persoonan näkökulmasta, ei muiden kokemuksista voi tietää mitään. Monet ovatkin hylänneet zombi-argumentin tämän takia: muiden (vaikkapa kaikkien maailmassa koskaan eläneiden ihmisten itseäni lukuun ottamatta) kokemusten kieltäminen on monelle liian kova hinta. Mutta hinta on maksettava, jos hyväksyy, että fenomenaalista kokemuksista saa tietoa vain ensimmäisen persoonan subjektiivisesta näkökulmasta.

Voidaan siis todeta, että modaaliargumentti osoittaa, että fysikaalisen kuvittelemisen ilman fenomenaalista on kuviteltavissa, siispä loogisesti mahdollista. Argumentista *ei seuraa*, että se olisi metafyyssisesti mahdollista (mutta ei myöskään, että se olisi metafyyssisesti mahdotonta). Van Gulick (2003, 329; 2004, 367) kutsuu

⁴³ Kaikki eivät tosin ole tästäkään samaa mieltä. Joidenkin moninaisen toteutuvuuden kritikoiden mielestä (esim. Kim 1992; Bickle 2008) neuraalinen plastisuus ei välttämättä kumoa identiteettiteoriaa, sillä plastisuus toimii kuitenkin jokseenkin systemaattisesti.

modaaliargumenttia *bumerangiargumentiksi*,⁴⁴ sillä se kurottuu subjektiiviselta alueelta objektiiviselle alueelle: faktoista kuinka *käsitteellistämme maailman* faktoihin millainen *maailma todella on*.

The distinctive feature of a boomerang argument is that it reaches across to the epistemic/cognitive/conceptual domain of facts about our representations of the world, and then swings back to reach a conclusion in the metaphysical/ontological/factual domain about the nature of reality itself (Van Gulick 2004, 367).

Tällainen kritiikki koskettaa kuitenkin vain Chalmersia, ei koko modaaliargumenttia. Modaaliargumentista ei tarvitse tehdä tällaista ontologista ”bumerangiversiota”, vaan sen avulla voi myös pelkästään osoittaa episteemisen kuilun kokemuksen ja fyysikaalisen rakenteen välissä. Episteeminen kuilukin on hyödyllinen lopputulos, sillä se osoittaa reduktiivisen fysikalismen vääräksi.

4. Tietoargumentti

Kolmas argumentti fysikalismia vastaan on *tietoargumentti* (*knowledge argument*). Tietoargumentin mukaan tietoisuuteen kuuluu jotain, joka ei ole deduktiivisesti johdettavissa fyysikaalisista faktoista. Tämän voi päätellä siitä, että joku voisi *tietää* kaikki fyysikaaliset faktat, mutta hän ei silti tietäisi kaikkia tietoisuuteen liittyviä faktoja. Kuuluisimman version tietoargumentista on esittänyt Frank Jackson vuonna 1982 ilmestyneessä artikkelissaan ”Epiphenomenal Qualia” (hän tarkensi argumenttia vielä ”What Mary Didn’t Know” -artikkelissa vuonna 1986).⁴⁵ Tietoargumentin taustalla on samanlainen episteeminen kuilu kuin kahdessa ensimmäisessäkin argumentissa: fyysikaalinen teoria ei selitä fenomenalisia faktoja. Tämänkin argumentin voi jättää episteemiselle tasolle, ja väittää että fenomenaliset

⁴⁴ Myös tietoargumentti ja selityksellisen argumentin ontologinen versio ovat bumerangiargumentteja. Van Gulick (2004, 367) paljastaa, että osasyynä nimitykseen on tällaisten argumenttien kuuluisimpien esittäjien kansallisuus: sekä David Chalmers että Frank Jackson ovat kumpikin australialaisia.

⁴⁵ Samansuuntaisia argumentteja on esitetty kuitenkin ennen Jacksoniakin, mm. C. D. Broadin (1925) matemaattinen arkkienkeli, joka ei tunnista ammoniakin hajua, vaikka tietääkin sen fyysikaalisen kuvauksen, ja B. A. Farrellin (1950) marsilainen supertiedemies, joka ei kaikesta tietämyksestään huolimatta tiedä miltä tuntuu olla ihminen. Jacksonin muotoilema tietoargumentti on kuitenkin niitä (kuten kaikkia muitakin lukemiani argumentin versioita) selitysvoimaisempi, sillä värit ovat selkeämpi esimerkki kuin muut aistihavainnot.

faktat kuitenkin *ovat* fysikaalisia, vaikka fysiikka ei niitä voikaan selittää. Jacksonin muotoilemana tietoargumentti oli kuitenkin alun perin tarkoitettu ontologiseksi argumentiksi, jossa episteemisestä kuilusta seuraa ontologinen kuilu.

Jackson esittää argumenttinsa kahden ajatuskokeen muodossa (1982, 128-130). Ensimmäinen esimerkki kertoo Fredistä, jolla on poikkeuksellisen kehittynyt värinäkö. Siinä missä muut näkevät vain punaista, erottaa Fred kaksi eri väriä. Vaikka meillä olisi *kaikki fysikaalinen tieto* Fredistä, emme kuitenkaan tietäisi, millaista on havaita ylimääräinen väri, eli meillä ei olisi *kaikkea tietoa* Fredistä. Siispä fysikalismi jättää jotain ulkopuolelleen. Toinen – ja kuuluisin – esimerkki kertoo neurotieteilijä Marysta, joka on viettänyt koko elämänsä mustavalkoisessa huoneessa, ja mustavalkoisten kirjojen ja television avulla saavuttanut parhaan mahdollisen neurotieteellisen sivistyksen. Etenkin hän on perehtynyt juuri näköhavainnon fysiologiaan. Ei ole yhtäkään fysikaalista faktaa, jota Mary ei tiedä värihavainnoista. Mutta mitä tapahtuu, kun Mary vapautetaan huoneesta? Oppiiko hän jotain uutta nähdessään ensimmäisen kerran punaista?

Tämä kysymys muodostaa siis koko tietoargumentin. Jackson vastasi kysymykseensä (vuoden 1998 mielenmuutokseen asti)⁴⁶ myöntävästi: kyllä, Mary oppii jotain uutta maailmasta ja värien havaitsemisesta, eli Maryn fysikaalinen tieto ei ollut kaikki tieto, ja Jacksonin mukaan tästä seurasi, että fysikalismi on väärässä. Muotoilen tietoargumentin tähän vielä mahdollisimman selkeästi:

Premissi 1. Marylla on *kaikki fysikaalinen tieto* värien havaitsemisesta ennen vapautumistaan.

Premissi 2. Marylla ei ole *kaikkea tietoa* värien havaitsemisesta ennen vapautumistaan.

Johtopäätös 1. Värien havaitsemisesta on olemassa tietoa, joka ei ole fysikalistista tietoa.

Johtopäätös 2. Fysikalismi on väärässä.

⁴⁶ Jacksonin siirtymistä vastustajien leiriin on kutsuttu yhdeksi merkittävimmistä nykyfilosofian täyskäännöksistä. Mielenmuutosta uutisoitiin mm. otsikolla ”Frank Jackson Changes His Mind”: *“Frank Jackson is now a physicalist. Yes, the same Frank Jackson who used to call himself a ‘qualia freak’ now claims that qualia [...] can be given a complete physical explanation. After more than a decade of defending his anti-physicalist views, Jackson has come around to the other side: physicalism, he now claims, is true.”* (Montero, 1999.)

4.1. Jacksonin tietoargumentin fysikalismi

Jackson itse ei tarkenna, minkä muotoista fysikalismia vastaan hän on argumenttinsa suunnannut. Hän sanoo vain, että fysikalismi on teesi, jonka mukaan kaikki (oikea) informaatio on fysikaalista informaatiota (Jackson 1982, 127). Jackson siis postuloi melko vapaasti jonkinlaisen fysikaaliseen informaatioon perustuvan fysikalismin, jota sitten vastustaa. Hänen mukaansa ”fysikaalinen informaatio” on sitä, mitä fysiikka, kemia, biologia ja niistä seuraavat tieteet (kuten neurofysiologia) käsittelevät.

For example, if a medical scientist tells me enough about the processes that go on in my nervous system, and about how they relate to happenings in the world around me [...] he or she tells me – if I am clever enough to fit it together appropriately – about what is often called the functional role of those states in me [...]. This information, and its kin, I also label ‘physical’. (Jackson 1982, 127).

Jacksonin esittämänä fysikalismi on siis episteeminen kanta. Kumpikaan tietoargumentin esimerkeistä ei vastusta suoraan identiteettiteoriaa eikä funktionalismia. Koska Jackson ei esitä väitteitä fysikalistisesta ontologiasta, jää fysikalistille tuttu pelastusreitti: vaikka fysikaalinen informaatio (tietoargumentin *tieto*) ei olisikaan palautettavissa fysikaaliseen, on kaikki kuitenkin ontologisesti fysikaalista. Fredin värihavainnot ovat erilaisia kuin muilla, mutta hänen aivorakenteensakin on erilainen.

We find out perhaps that Fred’s cones respond differentially to certain light waves in the red section of the spectrum that make no difference to ours (or perhaps he has an extra cone) and that this leads in Fred to a wider range of those brain states responsible for visual discriminatory behavior (Jackson 1982, 129).

Esimerkin mukaan jos muiden ihmisten aivorakenteen muuttaisi (jollain edistyksellisellä leikkauksella) samanlaiseksi kuin Fredin aivorakenne on, niin muutkin näkisivät saman värin. Tämä on täysin yhteensopivaa sekä identiteettiteorian että funktionalismin kanssa. (Tietoargumentti koskee vain sitä, että ennen leikkausta mikään määrä fysikaalista tietoa ei riittäisi kertomaan, millainen

ylimääräinen väri on.) Maryn tapaus on samanlainen (ainakin Jacksonin esittämän hypoteesin mukaan): kun Mary vapautetaan huoneesta, myös hänen aivotilansa muuttuvat.

After she is let out, things change; and physicalism can happily admit that she learns this; after all, some physical things will change, for instance, *her brain states and their functional roles*⁴⁷ (Jackson 1986, 292).

Tässäkään ei ole mitään ongelmaa identiteettiteorian tai funktionalismin suhteen, vaan Maryn tapaus sujuu juuri kuten identiteettiteoreetikot ja funktionalistit odottavatkin. (Tietoargumentti taas koskee sitä, että vapautuessaan Mary huomaa, kuinka tietämätön hän on ollut muiden värikokemuksista, vaikka tiesikin kaiken muiden aivorakenteesta ja sen funktionaalisista rooleista.) Tietoargumentti ei siis puuttuvan- ja käänteisen kvalian argumenttien tavoin vastusta suoraan funktionalismia eikä moninaisen toteutuvuuden argumentin tapaan identiteettiteoriaa. Se muistuttaa pikemminkin selityksellistä argumenttia, sillä sen kohde on epistemologinen ”fysikaalinen informaatio”.

Kripkeläinen kysymys a posteriorin välttämättömyydestä liittyy fysikalismiin määrittelyyn olennaisella tavalla. Voiko fakta olla välttämätön ja a posteriori? Fysikalismien suhteen kysymys on siis siitä, pidetäänkö tietoisuuden ja fysikaalisen suhdetta – toisin sanoen psykofyysistä konditionaalia (*jos* kaikki fysikaaliset totuudet, *niin* kaikki psykologiset totuudet) apriorisena vai aposteriorisena. Jos hyväksytään fysikalismien supervenienssi-teoria, siitä seuraa, että jos fysikalismi on totta, niin psykofyysinen konditionaali on välttämättä totta. Mutta kuten Kripke (1980) on tunnetusti osoittanut: se että jokin on välttämätöntä, ei ole sama asia, kuin että se on a priori. Psykofyysinen konditionaali voisi siis olla välttämätön ja a posteriori: toisin sanoen tietoinen kokemus olisi välttämättä fysikaalista, mutta ei a priori johdettavissa fysikaalisesta. Mary ei siis olisi voinut tietää, miltä punainen näyttää, ennen punaisen kokemusta (a posteriori), mutta punaisen kokemus on kuitenkin *välttämättä* fysikaalinen.

⁴⁷ Kursivointi omani.

[A]lthough physicalists are committed to the experiential being necessitated by a rich enough physical account of our world – otherwise it would take more than the physical nature of our world to secure its experiential nature, contrary to physicalism – they are not committed to the experiential being a priori derivable from the physical (Jackson 2002, 423).

Jacksonin tietoargumentin postuloiman fysikalismin psykofyysinen konditionaali on kuitenkin sekä *apriorinen* että välttämätön. Jackson on myöhemminkin (2002, 423) vastustanut kantaa, jonka mukaan tietoisuuden ja fysikaalisen suhde voisi olla samanlainen välttämätön a posteriori -totuus kuin esimerkiksi veden ja H₂O:n suhde. Välttämätön apriorisuus on olennaista, koska tietoargumentti tarvitsee vastustajakseen juuri a priori -selitykseen perustuvan fysikalismin.

Jackson (2002, 424) myöntää, että välttämättömiä a posteriori -tapauksia on kyllä olemassa, mutta väittää, että psykofyysinen konditionaali ei kuulu niihin. Hänen mukaansa välttämättömän aposteriorisen väitteen voi johtaa a priori joistain muista kontingenteista väitteistä. Mutta jos fysikalismi on totta, ei ole muita kontingenteja väitteitä (fysikalismi sisältää jo kaiken). Jackson väittää siis, että jos fysikalismi on totta, niin psykofyysinen konditionaali ei ole pelkästään välttämätön, vaan myös a priori. Voidaankin siis kysyä: onko Jackson oikeassa, eli onko fysikalismin psykofyysinen konditionaali välttämätön a priori, vai voisiko se kuitenkin olla välttämätön a posteriori? Jos fysikalismi todella vaatii Jacksonin esittämää välttämätöntä apriorisuutta (ollakseen fysikalismia ylipäätään), niin mielestäni on selvää, että tietoargumentti kumoaa fysikalismin.⁴⁸

Joko siis fysikalismi tai Jacksonin vaatimus apriorisuudesta on väärässä. (Tai sitten molemmat ovat väärässä.) Tarkastellaan ensin Jacksonin apriorisuus-väitettä. Sen mukaan fysikaalisen tiedon tulisi siis voida selittää a priori tietoinen kokemus (esim. kokemus punaisen havaitsemisesta). Kuinka tällainen olisi teoriassakaan mahdollista? Fysikaalisen kolmannen persoonan tiedon avulla pitäisi siis pystyä selittämään subjektiivinen kokemus ennen subjektiivista kokemusta. Jotkut fysikalistit toki pitävät tällaista täysin mahdollisena (esim. Dennett 1991), mutta

⁴⁸ Jackson (2002) on tästä eri mieltä: 2000-luvulla hän on väittänyt, että tietoargumentti on joka tapauksessa väärässä, vaikka psykofyysinen konditionaali onkin välttämätön a priori.

omasta mielestäni tämä on kuitenkin liikaa vaadittu fysikalismilta. Fysikalistinen oppi (ainakin fysikalistinen ontologia) voi toimia ilman apriorisuuden vaatimustakin. Siispä Jacksonin väite on väärässä, ja psykofyysinen konditionaali voi olla välttämätön ja a posteriori fysikalismin kanssa yhteensopivasti. Entä voisiko fysikalismi olla väärässä? Koska (kuten äsken osoitin) Jacksonin apriorisuuden vaatimus ei päde, ei tästä päättelystä seuraa, onko fysikalismi oikeassa vai väärässä. Tietoargumentti osoittaa vain, että Jacksonin esittämä apriorisuuteen perustuva fysikalismi on väärässä, mutta aposteriorisuuden sallivan fysikalismin selviäminen jää Jacksonin muotoilemassa argumentissa avoimeksi.

Jacksonin versio tietoargumentista on altis kritiikille myös sen takia, että hän väittää, että fysikalismin ulottumattomissa olevat faktat, kvaliat, ovat *epifenomenaalisia*. Siispä Maryn nähdessä ensi kerran punaista, hänen aivotilansa ja niiden funktionaaliset roolit kyllä muuttuvat, mutta syyt tähän ovat puhtaasti fysikaalisia. Värikokemus sisältää fysikaalisen puolen lisäksi myös kvalian, mutta sillä ei ole vaikutusta mihinkään, ainakaan mihinkään fysikaaliseen (Jackson 1982, 133). Fysikalistin on helppo sanoa, että jos kvalia ei kerran vaikuta mitenkään mihinkään, on fysikalismi totta, sillä ei ole tarpeen käsitellä tällaisia epäilmiöitä. Mutta kvalian määritelmäänhän juuri kuuluu, että ne ovat jotain: on *jonkinlaista* kokea niitä. Juuri tämä jonkinlaisuus (eli fenomenalaisuus tai *kvalitatiivisuus*) on se, mitä termi *kvalia* on luotu ilmaisemaan. Ja koska sen kokeminen on jonkinlaista, on sillä jotain vaikutusta johonkin. Vain jonkinlaisuuden puute mahdollistaisi vaikutuksettomuuden. Mutta jos kokeminen olisi ei-minkäänlaista, niin se ei olisi kokemista lainkaan. Jos kvaliat todella olisivat epifenomenaalisia, niin me olisimme kaikki zombeja.⁴⁹

Jacksonin muotoilema tietoargumentti johtaa episteemisestä kuilusta ontologiseen kuiluun, mutta ilman perusteita. Jos Chalmersin perustelut (ks. kappaleet 2.1. ja 3.4.) tällaisen johtamisen oikeutukselle ovat päteviä, niitä voi toki soveltaa tähänkin tapaukseen. Jos tietoargumentti koskee vain epistemologiaa, voi se osoittaa korkeintaan fysikalismin kyvyttömyyden selittää tietoisuus. Jos selityksen puutteesta edelleen seuraa ontologinen kuilu, on fysikalismi pohjimmiltaan väärässä. On

⁴⁹ En käsittele enempää Jacksonin epifenomenalismia, sillä se ei ole olennaista tietoargumentin kannalta. Argumentti toimii ilman epifenomenalismiaakin.

tilanne sitten kumpi tahansa, on ensin tarkasteltava, onko tietoargumentti ylipäätään pätevä argumentti, sillä jos se on mahdollista kumota tai osoittaa väärin muodostetuksi, on selvää, ettei sen perusteella voi päätellä niin ontologisia kuin episteemisiäkään seikkoja fysikalismien ja tietoisuuden suhteesta. Tarkastelen siis seuraavaksi tapoja, joilla sitä on pyritty – ja joilla sitä on ylipäätään mahdollista pyrkiä – kumoamaan. Esitän kritiikkimahdollisuudet kysymyssarjan muodossa.⁵⁰ Kysymysketjun avulla on tarkoitus selvittää, osoittaako argumentti episteemisen kuilun (ja jos osoittaa, seuraako siitä ontologinen kuilu). Van Gulickin (2004, 384) sanoin: ”Choose your favorite. Take your pick.”

4.1.1. Kysymys 1: Oliko Marylla todella kaikki fysikaalinen tieto ennen vapautumistaan?

Kysymys kohdistuu argumentin ensimmäiseen premissiin: onko ylipäätään mahdollista, että Marylla olisi ollut kaikki fysikaalinen tieto värien havaitsemisesta ennen vapautumistaan, eli ennen ensimmäistä värihavaintoa? Argumentin oletuksen mukaan hänellä tietenkin oli kaikki fysikaalinen tieto, mutta onko tällaista oletusta todella mahdollista tehdä? Onko periaatteessakaan mahdollista tietää kaikkea *fysikaalista* punaisen havaitsemisesta, jos ei ole koskaan havainnut punaista?

Jos kaikki fysikaaliset faktat ovat tiedettävissä pelkästään mustavalkoisten lähteiden avulla, niin on mahdollista tehdä hypoteesi, jonka mukaan Marylla todella oli tuolloin kaikki fysikaalinen tieto. (Vaikkei kukaan todellinen ihminen tähän välttämättä yltaisikään, *hypoteesi* on silti mahdollinen.) Mutta jos kaikkien fysikaalisten faktojen tietämiseen tarvitaan tietoa myös *subjektiivisista* (fysikaalisista) faktoista, jotka ovat ymmärrettävissä vain jostain tietystä näkökulmasta, niin silloin Marylla ei olisi edes periaatteessa voinut olla kaikkea fysikaalista tietoa ennen vapautumistaan, koska hänellä ei ollut subjektiivista kokemustietoa värien havaitsemisesta. Van Gulickin (2004, 390–391) mukaan *subjektiiviset fysikaaliset faktat* ovat kyllä ontologisesti fysikaalisia, mutta kognitiivisesti saavutettavissa vain tietyn kokemuksellisen näkökulman kautta. Jos

⁵⁰ Alun perin tällaisen kysymyssarjan on esittänyt Robert Van Gulick (1993; 2004). Daniel Stoljar ja Yujin Nagasawa (2004) ovat rakentaneet oman versionsa sen pohjalta. Yhdistelen hieman näitä kahta sarjaa, joten tässä esittämäni kysymyssarja on siis peräisin näistä molemmista lähteistä, omilla painotuksillani muokattuna.

tällaisia faktoja on todella olemassa, Marylla *ei voinut olla* kaikkea fysikaalista tietoa ennen vapautumistaan.

Jos fysikalismi siis sallii subjektiivisia fysikaalisia faktoja, tietoargumentti kaatuu jo heti ensimmäisen premissin kohdalla. Tällaista kantaa näyttäisi tukevan esimerkiksi Nagelin (1974) argumentti. Jyrkän reduktiivisen fysikalismin kannattajat ovat tietenkin eri mieltä: heidän mukaansa on (ainakin periaatteessa) mahdollista oppia kaikki fysikaaliset faktat mustavalkoisista lähteistä, eivätkä he pidä ensimmäistä premissiä ongelmallisena. Tämän kannan edustajia ovat mm. Paul M. Churchland ja Daniel Dennett. Jos seuraa heidän linjaansa, eteen tulee lisää kysymyksiä.

4.1.2. Kysymys 2: Oppiiko Mary jotakin uutta?

Toinen kysymys nousee suoraan tietoargumentista: oppiiko Mary jotain uutta, kun hän näkee ensimmäisen kerran punaista? On selvää, että jos vastaus on kieltävä, tietoargumentti kaatuu, ja fysikalismi on (ainakin tältä osin) pelastettu. Kieltävän vastauksen ovat antaneet ainakin edellä mainitut eliminativistiset materialistit Churchland ja Dennett. He väittävät siis, ettei Mary opi mitään uutta väreistä, kun hän saa ensimmäisen värikokemuksensa: Maryllahan jo hypoteesin mukaan oli *kaikki fysikaalinen tieto*, mikä siis on fysikalisteille sama asia kuin *kaikki tieto*. Oppiminen ei ole tässä tilanteessa yksinkertaisesti mahdollista. Tällainen näkemys on mahdollista muodostaa, vaikka se onkin intuition vastainen: se ei suostu vastaamaan tietoargumentin esittämään kysymykseen, koska jo ensimmäinen premissi kertoo kaiken oleellisen Marysta. Toisen premissin kuvaama asiantila ei siis tämän näkemyksen mukaan ole mahdollinen, jos ensimmäinen premissi on asetettu, sillä fysikalisteille *kaikki fysikaalinen tieto = kaikki tieto*. Kanta on johdonmukainen reduktiivisen fysikalismin kannattajille, jos tarkoituksena on varjella fysikalismin oppia. Chalmersin sanoin voisi kuitenkin huomauttaa, että tällainen näkemys ”ei ota tietoisuutta vakavasti”.

Churchland (1985) ja Dennett (1991) ovat jotakuinkin samoilla linjoilla Marysta: koska hän kerran tietää kaiken värihavainnon neurologiasta, niin hän tietää (mustavalkoisessa huoneessaankin) erinomaisesti, miltä vaikkapa punainen näyttää.

Hänhän pystyy johtamaan sen suoraan neurologisista tiedoistaan. Churchlandin (1985, 25–26) mukaan Mary voisi oppia käsitteellistämään havaintojaan siten, että hänen näköhavaintonsa eivät olisi enää pelkästään ”mustan kokemista”, ”valkoisen kokemista” ja ”harmaan kokemista”, vaan hän yhdistäisi jokaisen kokemuksen tiettyyn valoallon taajuuteen. Eikä tämä päde vain Maryn jo havaitsemiin harmaan eri sävyihin, vaan hän voisi kuvitella neurotieteellisten käsitteidensä avulla muutkin värit, vaikka ei olisi koskaan nähnyt niitä.

If Mary has the relevant neuroscientific concepts for the sensational states at issue (viz., sensations-of-*red*), but has never yet been *in* those states, she may well be able to imagine being in the relevant cortical state, and imagine it with substantial success, even in advance of receiving external stimuli that would actually produce it (Churchland 1985, 26).

Churchlandin mukaan Mary osaisi vapauduttuaan tunnistaa esimerkiksi punaisen värin, jos hän vain olisi selvittänyt sen taajuuden ja päätellyt siitä, miltä punainen näyttää – toisin sanoen hän olisi kuvitellut punaisuuden. Mielestäni tällaisessa ajattelussa on kuitenkin hyvin ongelmallista se, että Mary ei voisi saada tietää, miltä punainen näyttää *kokematta* miltä punainen näyttää. Siis jos hän pystyisi neurotieteellisten kykyjensä avulla päättämään, miltä tuntuu kokea punaista, niin *hän kokisi punaista*. Tietoargumentin kannalta on aivan samantekevää, vapautuuko Mary huoneesta, annetaanko hänelle red delicious -omena, vai *kuvitteleeko* hän punaista. Kuten Kripke (1980) on huomauttanut, ei tietoisesta kokemuksesta voi erottaa *tietoisesta kokemuksen tuntemusta*. ”If any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon *is* pain” (Kripke 1980, 153). Siispä Churchland tulee vain keksineeksi Maryn tarinalle erilaisen lopun, joka ei kuitenkaan vaikuta mitenkään argumentin sisältöön.

Dennettin (1991, 399–400) mukaan Maryn tapaus on seuraavanlainen: Eräänä päivänä Maryn vangitsijat päättivät, että Maryn on aika nähdä värejä, ja he maalasivat ensimmäistä värikokemusta varten banaanin siniseksi. Mary katsoi banaania, ja huomasi heti, että häntä oli huijattu – banaanithan ovat keltaisia eivätkä sinisiä. Vangitsijat olivat tyrmistyneitä: kuinka Mary pystyi tähän? ”Yksinkertaista”, Mary vastasi, ja muistutti vangitsijoitaan siitä, että hän tiesi *kaiken* tiedettävissä

olevan värihavaintojen fysiikasta. Ennen banaanin näyttämistä Mary oli kirjoittanut itselleen tarkan kuvauksen, millaisen vaikutuksen keltainen objekti tai sininen objekti (tai vihreä objekti jne.) saisi aikaan hänen hermojärjestelmässään. Hän tiesi siis tarkalleen, millaisia ajatuksia hänellä tulisi olemaan.

Dennettin mukaan on selvää, että Mary tietäisi ilman värikokemustakin, miltä punainen näyttää, jos hän tietäisi kaiken fysikaalisen – siis ehdottoman *kaiken* fysikaalisen. Mutta hänen mukaansa on vaikeaa kuvitella tilanne, jossa henkilö todella tietäisi kaiken fysikaalisen. Tämä vaikeus aiheuttaa Dennettin (1991, 400) mukaan koko sekaannuksen, eli intuition siitä, että Mary oppisi jotain uutta saadessaan ensimmäisen värikokemuksen. Dennett siis väittää, että uuden oppimisen intuitio syntyy siitä, että koska on niin vaikeaa kuvitella henkilöllä olevan *kaikkea* fysikaalista tietoa, niin tullaan kuvitelleeksi hänellä sen sijaan olevan *hyvin suuri määrä* fysikaalista tietoa. Ja silloin uuden oppimiselle jää mahdollisuus.

Omasta mielestäni on taas *hyvin vaikeaa* kuvitella, kuinka mustavalkoisista fysikaalisista faktoista voisi päätellä punaisen aistikokemuksen – oli faktoja sitten kuinka paljon tahansa. Se, että tiedossa ovat *kaikki* fysikaaliset faktat, ei paranna tilannetta lainkaan. Howard Robinson (1993, 70–71) muistuttaa kritiikissään Dennettiä kohtaan, että Mary ei ole kaikkietävä: hänellä on vain kaikki fysikaalinen (so. tieteellinen) tieto. Mary ei siis tunne ennalta jokaista maailmassa olevaa fysikaalista partikulaaria, eikä hän tiedä jokaista aivotapahtumaansa jokaisena ajanhetkenä. Hän voi kyllä kirjoittaa itselleen ylös, millaisen vaikutuksen tietynvärinen objekti saa aikaan hänen hermojärjestelmässään, mutta katsoessaan jotakin objektia, hän ei voi tietää, onko sillä juuri tämä kyseinen vaikutus. Mary ei voi tietää, onko *tämän* objektin *tällaiselta* näyttäminen punaisen tai sinisen jne. vaikutelma.

Siispä voidaan todeta, että Dennettin ja Churchlandin näkemys siitä, että Mary voisi fysikaalisen tietonsa avulla päätellä, miltä värit näyttävät, on yksinkertaisesti väärä. Useimmat keskusteluun osallistuneet filosofit ovat kuitenkin antaneet myöntävän vastauksen tähän kysymykseen: on selvää, että Mary oppii jotain uutta vapautuessaan. Oleellinen kysymys onkin, että mitä tai millaista tuo jokin on. Siitä

on esitetty hyvin monenlaisia näkemyksiä, ja seuraavat viisi kysymystä (3–7) käsittelevät juuri näitä erilaisia tapoja vastata myöntävästi kysymykseen kaksi.

4.1.3. Kysymys 3: Onko oppiminen faktuaalista vai ei-faktuaalista?

Jos hyväksytään, että Mary oppii jotakin uutta vapautuessaan, on vielä selvitettävä, millainen asia on kyseessä. Karkein jaottelu on varmaankin erotettavissa faktuaalisen ja ei-faktuaalisen (tai propositionaalisen ja ei-propositionaalisen) tiedon välillä⁵¹. Gilbert Ryle (1949) on esittänyt, että ”tieto että” (*knowledge that*) ja ”tieto miten” (*knowledge how*) ovat toisistaan erillisiä tiedon lajeja.

Learning *how* or improving in ability is not like learning *that* or acquiring information. Truths can be imparted, procedures can only be inculcated, and while inculcation is a gradual process, imparting is relatively sudden. It makes sense to ask at what moment someone became apprised of a truth, but not to ask at what moment someone acquired a skill. (Ryle 1949, 58.)

Tietämisen voi korvata *oppimisella* Rylen erottelun muuttumatta. Voi siis kysyä, millä hetkellä joku *oppi* faktan, mutta ei millä hetkellä joku *oppi* taidon. Erottelua voi havainnollistaa seuraavalla esimerkillä:

1. Patricia tietää/oppii, *että* Paul ei ole matolla.
2. Patricia tietää/oppii, *miten* huijataan veroviranomaisia.

Toisen lauseen *miten*-tieto on Rylen mukaan samanlaista kuin tieto *kuka*, *mikä*, *milloin*, *missä* jne., toisin sanoen ei-faktuaalista tietoa/oppimista, kun taas ensimmäisen lauseen tieto/oppiminen on faktuaalista. Kummanlaista tietoa Mary saa? Jos uusi tieto ei ole faktuaalista, tietoargumentti kumoutuu, sillä argumentti vaatii toimiakseen sen, että Mary oppii aidosti uuden proposition eli saa uutta faktuaalista ja totuusarvollista tietoa.

⁵¹ Käytän tässä termejä ”fakta” ja ”faktuaalinen” synonyymisesti termien ”propositio” ja ”propositionaalinen” kanssa.

Tiedon lajien erottelusta tosin herää kysymys, ovatko lajit keskenään riittävän erilaisia, jotta toisen tavan tieto on tietoa sanan varsinaisessa merkityksessä, ja toisen tavan ei. Ja onko Maryn tapauksen kannalta väliä, kummanlaista tietoa hän saa? Voisi nimittäin väittää, että tietoargumentin kannalta minkälainen tahansa uusi tieto riittää. Tietenkin jos ei-faktuaalinen ”tieto” todella on jonkinlaista ”epätietoa”, Maryn oppiminen osoittautuu näennäiseksi. Lisäksi voi tietenkin kysyä, onko näin jyrkkää erottelua faktuaalisen ja ei-faktuaalisen tiedon välillä *ylipäättään* mahdollista tehdä. Olipa erottelu sitten todellisuutta vastaava tai ei, ainakin monet filosofit ovat perustaneet näkemyksensä tietoargumentista sen varaan. Seuraavat kaksi kysymystä käsittelevät ei-faktuaalisia vastauksia tähän kolmanteen kysymykseen, ja kuudes ja seitsemäs kysymys taas faktuaalisia ratkaisuja.

4.1.4. Kysymys 4: (Jos ei-faktuaalista, niin) saako hän vain uuden taidon?

On mahdollista esittää, että Mary oppii kyllä jotakin uutta vapautuessaan, mutta kyseessä ei ole uusi fakta. Yksi mahdollisuus on väittää, että Maryn oppima uusi asia on vain uusi *taito* tai *kyky*. Kutsun näkemystä tässä *taitotieto-hypoteesiksi*.⁵² Tämä on kenties tietoargumentin kuuluisin (joskaan ei onnistunein) vastalause, jonka on esittänyt alun perin David Lewis (1988) ja Lawrence Nemirow (1990). Lewis (1988) esittää, että oppiessamme millainen uusi kokemus on (esimerkiksi uusi värikokemus), saamme uusia kykyjä *muistaa*, *kuvitella* ja *tunnistaa*. Mary oppii siis esimerkiksi punaisen tunnistamisen taidon. Hän ei opi mitään uutta propositionaalista faktaa, sillä kaikki ne hänellä oli jo tiedossa entuudestaan.

Know-how is ability. But of course some aspects of ability are in no sense knowledge: strength, sufficient funds. Other aspects of ability are, purely and simply, a matter of information. If you want to know how to open the combination lock on the bank vault, information is all you need. It remains that there are aspects of ability that do *not* consist simply of possession of information, and that we *do* call knowledge. The ability hypothesis holds that knowing what an experience is like is that sort of knowledge. (Lewis 1988, 100.)

⁵² (Vrt. *ability hypothesis, know-how*.) Taitotieto on mielestäni tässä tapauksessa kuvaavampi nimitys kuin tietotaito, sillä se korostaa taitoa enemmän kuin tietoa.

Hypoteesi perustuu siis suoraan Rylen erottelemiin faktuaaliseen ja ei-faktuaaliseen tietoon: ei-faktuaalinen ”tieto” ei ole tietoa sanan varsinaisessa merkityksessä, vaan pelkkä kyky. Useissa tapauksissa eri tavat vain ovat sekoittuneina keskenään (kuten vaikkapa kirjoitustaidossa – se vaatii niin faktuaalista informaatiota kuin ei-faktuaalista kykyä), joten niitä saatetaan erehtyä pitämään puhtaana tietona. Taitotieto-hypoteesin mukaan Maryn tapaus on seuraavanlainen: Marylla oli kaikki fysikaalinen tieto ennen vapautumistaan, ja koska fysikalismi on totta, niin Marylla oli siis *kaikki* tieto ennen vapautumistaan. Vapautuessaan hän kuitenkin oppii jotain uutta, mutta kyseessä ei ole uusi tieto, vaan vain uusi taito. Taito saattaa vaikuttaa tiedon lajilta, koska se esiintyy usein tiedon kanssa, mutta sitä se ei kuitenkaan ole. Maryn tiedollinen asema on aivan sama ennen vapautumista ja sen jälkeen.

Taitotieto-hypoteesi on kuitenkin monella tavalla ongelmallinen. Vaikuttaisi kyllä selvältä, että Mary saa jonkinlaisia uusia taitoja nähdessään ensimmäisen kerran punaista – vaan ei kuitenkaan lainkaan selvältä, ovatko nämä taidot *ainoa* asia, jonka hän oppii. Monet (niin argumentin kannattajat kuin vastustajatkin) ovat kritisoineet taitotieto-hypoteesia, ja väittäneet, että Mary oppii muutakin kuin vain pelkkiä uusia taitoja.⁵³

Lewisin (1988) väite, että *tieto miten on kyky* (”*know-how is ability*”), on jo sinänsä kyseenalainen. Yksinkertainen vastaesimerkki osoittaa, että näin ei ole: kitaransoitonopettaja saattaa *tietää miten* soitetaan hyvin, mutta hänellä ei kuitenkaan välttämättä ole *kykyä* soittaa hyvin. Opettajan *kyky soittaa hyvin* saattaa myös parantua harjoittelemalla, vaikka *tieto miten soitetaan hyvin* säilyykin samana. Tieto miten (taitotieto) ei siis redusoidu kykyyn. (Alter 2001, 232.)

Taitotieto-hypoteesiä voi kritisoida myös väittämällä, että oikeastaan Mary ei saa *edes* niitä taitoja, joita Lewis väittää hänen saavan. Kun Mary näkee ruusun, hän näkee kyllä sen punaisuuden, mutta hänen on mahdotonta *muistaa, kuvitella* tai *tunnistaa* täsmälleen samaa punaisen sävyä myöhemmin – onhan hän tavallinen

⁵³ Jo Jackson itse vastusti taitotieto-hypoteesia artikkelissaan ”What Mary Didn’t Know” (1986, 294–295). Jacksonkin myöntää, että Mary varmasti saa myös uusia kykyjä vapautuessaan, mutta kykyjen lisäksi hän saa myös uutta tietoa.

ihminen, jolla on tavallinen näkökyky. Sen sijaan tuijottaessaan ruusua, hän *tietää* tarkalleen *millaista* (*what it is like*) on kokea juuri sen punaisen sävyn näkeminen. Marylla on siis tieto ”millaista jokin on”, mutta ei kykyä. (Tye 2000, 151–152.)

Perhaps knowing what it is like should be identified not with the cluster of abilities Lewis cites – for they may all be lacking while knowing what it is like is present – but rather with the more basic ability to apply an indexical concept to the phenomenal character of her experience via introspection. (Tye 2000, 153.)

Tye (2000, 153) antaa myös päinvastaisen esimerkin. Mary voi esimerkiksi olla uppoutunut muihin ajatuksiin nähdessään ruusun, joten hän ei välttämättä kiinnitä siihen huomiota. Hänen silmänsä ovat auki, ja punainen ruusu on hänen edessään. Mutta koska hän on hajamielinen, hän ei muodosta ruusun näkemisestä mitään käsitystä, eli hän ei siis tässä tilanteessa *tiedä millaista* on ruusun näkemisen kokeminen. *Kyky* (suuntautua indeksikaalisesti tietynhetkiseen kokemukseen introspektion avulla – ”tämä punaisen sävy, jonka havaitsen tällä hetkellä”) sen sijaan on tallella. Tässä tapauksessa Marylla on siis kyky, mutta ei tietoa ”millaista jokin on”.

Yksi kritiikin muoto taitotieto-hypoteesille on myös ns. sokeanäkö. Henkilö, joka ei ole koskaan nähnyt värejä, saattaa pystyä silti tunnistamaan eri värejä toisistaan, vaikka hänellä ei olekaan tietoisia värikokemuksia. Tämä vaatii tilanteen, jossa henkilön silmät ovat kunnossa, mutta näköaivokuori (tai silmän ja aivokuoren väliset yhteydet) ovat vaurioituneet. Tällaisella henkilöllä voi olla sama *kyky* kuin normaalisti näkevilläkin, vaikka häneltä puuttuu *tieto* siitä, millaista kokeminen on. (Bigelow & Pargetter 1990, 193.)⁵⁴ Tieto siitä millaista jokin on, ei siis ole sama asia kuin kyky. Tässä tulee esiin tietoisuuden käsitteen monimerkityksisyys: taitotietohypoteesin kannattajat käsittävät tietoisuuden *psykologisena pääsy tietoisuutena*,

⁵⁴ Esimerkkejä kyvyn ja fenomenaalisen tiedon eroista on lukuisia muitakin, esimerkiksi ns. *split-brain* –potilaat, joiden aivopuoliskojen välillä tieto ei siirry normaalisti, ja tämän takia he osaavat kuvailla sanallisesti, mitä (kielellisesti dominantin) vasemman aivopuoliskon näkökentässä on, mutta osoittaa vain motorisesti (sormella/kädellä jne.), mitä oikean aivopuoliskon näkökentässä on (Kalat 2004, 423). Tapaukset eivät ole kuitenkaan yksiselitteisiä, sillä vaikka henkilö väittäisikin (sanallisesti), ettei hänellä ole visuaalisia kokemuksia, niin se ei välttämättä todista, että hänellä ei niitä todellisuudessa ole.

vaikka tietoargumentissa on kyse *fenomenaalista tietoisuudesta*. Kritiikki ei siis vastaa oikeaan kysymykseen.

Nämä esimerkit osoittavat, että *tieto millaista jokin on* ja *kyky* eivät ole sama asia, sillä Marylla voi olla tieto olematta kykyä tai kyky olematta tietoa. Tieto ”millaista jokin on” voi toki olla (myös) joidenkin kykyjen hallitsemista. Mutta se ei ole välttämättä ja aina sitä, joten tieto ”millaista jokin on” ja kyky eivät ole identtisiä. Voisi siis sanoa, että Mary voi hyvinkin saada uusia kykyjä vapautuessaan, mutta uudet kyvyt johtuvat uudesta tiedosta – ja uusi tieto on enemmän kuin pelkät kyvyt. Taitotieto-hypoteesi on siis väärässä.

4.1.5. Kysymys 5: (Jos ei-faktuaalista, niin) saako hän vain tuttuustietoa?

Toinen mahdollinen selitys sille, mitä ei-faktuaalinen tieto on, on tuttuus-hypoteesi (*acquaintance hypothesis*). Tuttuus-hypoteesin mukaan vapautuessaan huoneesta Mary ei saa uutta tietoa, vaan vain *tulee tutuksi* uuden ominaisuuden kanssa. Tutuksi tuleminen punaisuuden ominaisuuden kanssa on sitä, että tietää, mitä punainen on. Taitotieto-hypoteesissa ja tuttuus-hypoteesissa on jokseenkin samanlaiset lähtökohdat, mutta niissä on myös eronsa.⁵⁵ Tuttuus-hypoteesin kannattajien mukaan *tuttuus* on oma tietämisen lajinsa:

Acquaintance with an experience does not require having either information or abilities. Acquaintance constitutes a third category of knowledge, irreducible to factual knowledge or knowing how. (Conee 1994, 197–198.)

Siispä Mary ei opi vapautuessaan uutta faktaa tai saa uutta kykyä, vaan hän tulee vain tutuksi uuden fenomenaalisen ominaisuuden kanssa, eli on uudeltaisessa kognitiivisessa (tuttuus-)suhteessa sen kanssa. Bigelowin ja Pargetterin (1990) mukaan Mary saa vapautuessaan tietoonsa saman asian, jonka hän tiesi jo aiemminkin, mutta nyt vain uudella tavalla.

⁵⁵ Taitotieto-hypoteesin kannattaja David Lewis (1988) esimerkiksi kritisoi tuttuus-hypoteesia. Earl Conee (1994) taas kannattaa tuttuus-hypoteesia, mutta kritisoi taitotieto-hypoteesia.

Before Mary had color experiences, she was acquainted with all the relevant structures which there are in the world. After having color experiences, she is acquainted in a new way with some of those structures. Her new knowledge state differs from the old ones, not at the level of truth-conditions, nor at the level of structures, but only at the third significant level, that of her modes of acquaintance with those structures. (Bigelow & Pargetter 1990, 188.)

Tuttuus-hypoteesin kannattajien mukaan siis fenomenalisista kokemuksista ei voi saada kaikkea tietoa kokematta niitä (koska tuttuus-tapa puuttuu), mutta silti on mahdollista kannattaa näkemystä, että fenomenaliset kokemukset palautuvat fysikaaliseen.

Tuttuus-hypoteesi kohtaa samat ongelmat kuin taitotieto-hypoteesikin. Ei ole selvää, että tuttuus on ainoa asia, jonka Mary oppii. Ongelmallista tässäkin kannassa on myös se, että jako faktuaaliseen ja ei-faktuaaliseen tietoon ei välttämättä ole aukoton. Tuttuus-hypoteesissa on lisäksi vielä se ongelma, että *tuttuus/tuttuustieto* on käsitteenä hyvin monimerkityksinen. Joskus sitä käytetään viitattaessa tietentyypiseen tietoon, joskus taas viitattaessa kokemukseen. Jos tuttuustieto tulkitaan kokemukseksi, niin hypoteesin mukaan saadessaan uutta tuttuustietoa Mary saa siis uuden kokemuksen – mikä on sanomattakin selvää. Tuttuustieto kokemuksena on siis pelkkä truismi. (Jos se taas tulkitaan tiedoksi, niin Mary saa uutta tietoa, joka on – tietoa.) (Stoljar & Nagasawa 2004, 18.)

Sekä tuttuustieto että taitotieto perustuvat vahvasti Rylen erotteluun tiedon lajeista. Jos erottelu hylätään, putoaa molemmilta hypoteeseilta pohja. Erottelua voidaan kritisoida väittämällä, että Rylen erottamat tiedon lajit eivät loppujen lopuksi olekaan niin erilaisia. Ei-propositionaalista ”tietoa miten” pidetään usein häilyvämpänä ja epävarmempana tiedon lähteenä kuin propositionaalista ”tietoa että”. Ei-propositionaalisen tiedon häilyvyyttä voi havainnollistaa seuraavan esimerkin avulla:

Patricia on aloittanut uuden harrastuksen: hän käy yllä mainitun kitaranopettajan pitämällä klassisen kitaransoiton tunneilla. Eräänä iltana harjoitellessaan seuraavan päivän tuntia varten hän tulee Villa-Loboksen ja Schubertin lomassa tapailleeksi säveliä, jotka muodostavat Slayerin ”Raining Bloodin”. Hänellä on siis kyky soittaa

kyseinen kappale, vaikka hän ei tiedä, miten se soitetaan. Seuraavana päivänä kokeilunhaluinen kitaraopettaja kysyy Patriciaalta, osaako hän soittaa Slayerin ”Raining Bloodin”. Patricia kieltää ehdottomasti. Ei ole kuitenkaan selvää, että Patricia ei osaa soittaa sitä (vaikka hän olikin täysin rehellinen).

Torin Alter (2001, 232) kuitenkin huomauttaa, ettei propositionaalinenkaan tieto ole välttämättä yhtään selkeämpää. Seuraako tiedosta, että Cicero oli roomalainen, tieto että Tullius oli roomalainen? Joku joka tietää, että Cicero oli roomalainen, saattaa rehellisesti kieltää tietävänsä, että Tullius oli roomalainen, vaikka kysymyksessä on todellisuudessa sama henkilö. Mutta tästä ei kuitenkaan yksiselitteisesti seuraa, että kyseinen henkilö ei tietäisi jälkimmäistä faktaa. Tässä mielessä *tieto miten* (soitetaan Slayerin ”Raining Blood”) ja *tieto että* (Tullius oli roomalainen) toimivat samalla tavalla. Siispä kysymys siitä, onko ei-propositionaalinen tieto liian epätarkkaa täyttääkseen ”oikealle tiedolle” vaadittavat kriteerit, on hyvin pitkälti sama kysymys kuin kysymys siitä, onko propositionaalinen tieto siihen liian epätarkkaa.

Siispä jako propositionaaliseen ja ei-propositionaaliseen tietoon ei ole välttämättä aukoton: ainakin sen pohjalta johdettuihin päätelmiin on suhtauduttava varauksella. Olen myös sitä mieltä, että taitotieto ja tuttuustieto eivät eroa toisistaan tietoargumentin kannalta olennaisessa mielessä. Puhuttiin sitten taidosta tai tuttuudesta, niin Maryn tiedollinen asema on näissä tapauksissa sama. (Riippumatta siitä, saako hän uutta tietoa vai ei: taito ja tuttuus joka tapauksessa joko selviävät tai kaatuvat yhdessä. Itse väitän, että kaatuvat.) Vaikka tällaisia erotteluita tiedonlajeista voisikin oikeutetusti tehdä, niin tässä tapauksessa tiedon lajilla ei ole merkitystä: argumentti toimii aivan samalla tavalla, oli Maryn saavuttama *uusi tieto* millaista tahansa.

4.1.6. Kysymys 6: (Jos faktuaalista, niin) saako hän tietoonsa uusia faktoja, vai vain vanhoja faktoja uudella tavalla ymmärrettynä?

Faktoja voi erotella eri tavoilla. Filosofien keskuudessa ei vallitse yksimielisyyttä siitä, mikä tulisi laskea selkeästi erilliseksi faktaksi. Jos hyväksytään, että Maryn saama uusi tieto on faktuaalista, voidaan siis edelleen kysyä, onko (faktuaalinen)

tieto aidosti uutta tietoa, vai sittenkin vain vanhaa (faktuaalista) tietoa, jota on vain erehdytty pitämään uutena. Useat filosofit ovat vedonneet fregeläiseen erotteluun ilmaisuuden mielestä (*Sinn*) ja merkityksestä (*Bedeutung*).⁵⁶ Ilmaisuuden *merkitys* (eli referenssi) voi olla sama, vaikka *mieli* on eri, johtuen esittämisen tavasta (*mode of presentation*). Seuraava esimerkki havainnollistaa fregeläistä eroa:

Patricia tietää, että Professor Chaos yrittää tuhota ihmiskunnan. Hän ei tiedä, että Butters Stotch yrittää tuhota ihmiskunnan. Jos Paul kertoo jälkimmäisen faktan hänelle, hän oppii jotain uutta. Tästä voisi päätellä, että Professor Chaos ei ole identtinen Butters Stotchin kanssa, mutta tämä ei pidä paikkaansa; kysymyksessä on todellisuudessa sama henkilö.

Uskomusten sisältö ei siis riitä määräämään niiden totuusarvoa, vaan vaaditaan myös tietty *esittämisen tapa* (*mode of presentation*). Tilanne voisi olla sama Maryn suhteen: ehkäpä hänkin saa vapautuessaan tietää jo tiedossaan olleen vanhan faktan vain uudella tavalla. Tässä tapauksessa tietoargumentti kaatuisi, sillä uusi fakta olisikin vanha fakta uudella tavalla ymmärrettynä.

Erottelu on sama kuin kappaleessa 3.4. esittelemässäni Chalmersin kaksiulotteisessa semantiikassa: Kripken a posteriorista -välttämättömyydestä johdettu erottelu primaaristen ja sekundaaristen intensioiden välillä. Primaarisella tavalla ymmärrettynä äskeisissä lauseissa on ero, koska ne eivät sisällä apriorista totuutta. Apriorisesti tarkasteltuna Professor Chaos ja Butters Stotch voisivat olla erimielisiä ihmiskunnan tuhoamisesta. Tämä on *episteemisesti mahdollista*. Primaarisessa tavassa lauseiden totuusarvoilla ei siis ole väliä – fregeläisittäin kyse on mielestä tai intensiosta, ei ekstensiosta. Sekundaarisella tavalla ymmärrettynä äskeiset lauseet taas ovat sama asia: empiirisen havainnoinnin jälkeen huomataan, että Professor Chaos ja Butters Stotch ovat identtisiä, joten jos toinen haluaa tuhota ihmiskunnan, niin välttämättä myös toinen haluaa tuhota ihmiskunnan. Totuusarvon kannalta mitä

⁵⁶ Esim. Tye (2000, 157) väittää, että Maryn oppima uusi tieto on samanlaista kuin tilanteessa, jossa joku oppii tietämään lauseen ”Cicero oli puhuja”, kun jo tiesi lauseen ”Tullius oli puhuja”. Vaikka lauseilla onkin sama totuusarvo, on silti mahdollista, että henkilö ei tiedä lauseista kuin toisen. Loar (1997, 222) taas vertaa Marya Maxiin, joka tietää, että pullossa on CH₃CH₂OH:ta, mutta ei tiedä, että pullossa on alkoholia. Max oppii uuden faktan saadessaan tietää, että pullossa oli alkoholia, mutta tämä ei tarkoita etteivät alkoholi ja CH₃CH₂OH olisi identtisiä.

sanotaan toisesta, on totta myös toisesta. Sekundaarinen tapa koskee Fregen termein lauseen *ekstensiota*, ja Chalmersin termein *sekundaarista intensiota*.⁵⁷

4.1.7. Kysymys 7: (Jos faktuaalista, niin) kumpi erottelutapa on oikea?

[T]he knowledge argument can allow us to deduce the falsity of physicalism from the incompleteness of Mary's propositional knowledge only if incompleteness entails her failing to know some proposition in the coarse-grained sense (Van Gulick 2004, 383).

Kumpi erottelun tapa sitten on oikea? Van Gulick (2004, 381) huomauttaa, että Maryn tapauksessa on oltava ainakin johdonmukainen. Väitteet ”Mary oppii uuden faktan” ja ”Jos fysikalismi ei kerro kaikkia faktoja, niin se on väärässä” tulee käsitellä samalla erottelutavalla. Jos Mary oppii uuden faktan vain primaarisella (hienojakoisella) tavalla, niin opitun asian totuusarvo ei ole uusi, vaan vain sisäinen käsitteellinen rakenne muuttuu. Siispä fysikalismi ei ole uhattuna. Sekundaarisella tavalla mielletynä uusi fakta sen sijaan näyttäisi osoittavan, että fysikalismi todella jättää jotain ulkopuolelleen. Tällä tavalla kyse ei ole pelkästä uudelleenkäsitteellistämisestä, vaan fakta on todella *uusi*.

However, supporters of the argument [...] claim that what Mary comes to know about the experience of seeing red involves not merely reconceptualization of old facts, but facts not previously known under any conceptualization. Given the highly abstract nature of the matters about which the two sides present their respective intuitions, it is difficult to see what might resolve the conflict. (Van Gulick 2004, 383.)

Jos erottelutapa on primaarinen, niin tietoargumentti siis kumoutuu saman tien, koska Mary oppii vain vanhan asian uudella tavalla. Mutta kuinka voisi vakuuttua, että erottelutapa todella on primaarinen? On jokseenkin epäselvää, voiko fenomenalisiin faktoihin soveltaa samanlaista erottelua kuin yllä esitettyihin

⁵⁷ Van Gulick (2004, 380) puhuu hienojakoisesta tavasta (*fine-grained mode*) primaarisen sijaan, ja karkeajakoisesta tavasta (*coarse-grained mode*) sekundaarisen sijaan. Karkeajakoisella tavalla ymmärrettynä vain faktojen totuusarvolla on merkitystä, joten tällä erottelutavalla ymmärrettynä esimerkkilauseet ovat identtisiä. Hienojakoisella tavalla lauseet sen sijaan eivät ole identtisiä, koska niiden *intensionaalinen rakenne* on eri, vaikka totuusarvo pysyykin samana.

esimerkkilauseisiin (Professor Chaos/Butters Stotch). Koska primaarisuus ei vaikuta olevan ainakaan selkeästi oikea tapa käsittää Maryn saama uusi fakta, on siis syytä tarkastella mitä sekundaarisesta tavasta seuraa.

4.1.8. Kysymys 8: Kumoaako Maryn sekundaarisella tavalla oppima uusi fakta fysikalismiin?

Äskeisestä kysymyksestä ei selvinnyt aukottomasti, kumpi erottelutapa on oikea. Se kuitenkin selvisi, että jos Mary oppii uuden faktan vain primaarisella tavalla, niin tietoargumentti ei kumoa fysikalismia. Fakta ei tässä tapauksessa ollutkaan todella uusi, vaan vain vanha fakta uudella tavalla ymmärrettynä. Mutta entä jos Mary oppiikin faktan sekundaarisella tavalla? Johtaako se ilman muuta fysikalismiin kumoutumiseen, vai voiko fysikalismi vielä senkin jälkeen pelastua? Jos hyväksyy, että Mary oppii faktan sekundaarisella tavalla, niin tietoargumentti osoittaa siis, että fysikaalisen tiedon lisäksi on muutakin tietoa. Mutta seuraako tästä, että fysikalismi on väärässä?⁵⁸ Entä jos fysikalismi ei olekaan väärässä, vaan Jackson ei vain ole muotoillut alun perinkään ”fysikalismia” oikein? Jacksonin vastustama fysikalismi on jyrkän reduktiivinen oppi, ja sellaisena tietoargumentilla kumottavissa. Mutta onko fysikalismi välttämättä käsitettävä Jacksonin esittämällä tavalla?

Tietoargumentti kysyy: oppiiko Mary mustavalkoisesta huoneesta vapauduttuaan jotain uutta? Se ei kysy, *minkälaista* tieto on. Vastaus kysymykseen on selvästi: kyllä oppii. Vaikka asiaa kuinka vatvoisi, muuhun vastaukseen ei voi päätyä. Mutta tästä ei tarkalleen ottaen tarvitse seurata, että fysikalismi (ontologisena oppina) on väärässä, vaan vain, että fysikalismi ei selitä tietoisuutta. Selityksen puutteesta johtuu, että Mary oppii vapautuessaan uuden faktan. Mutta Mary ei olisi periaatteessakaan voinut tietää tuota faktaa ennen vapautumistaan (koska hän ei ollut saanut subjektiivisella tavalla tietoa punaisen kokemisesta). Siispä koko argumentin perusta (Mary tiesi mustavalkoisessa huoneessa kaikki fysikaaliset faktat värien havaitsemisesta) on

⁵⁸ Se, mihin fysikalismiin (oletettu) kumoutuminen johtaisi, on taas toinen kysymys. Jacksonin mukaan fysikalismiin kumoutumisesta seuraa, että todellisuuteen kuuluu ei-fysikaalisia ominaisuuksia: hän kannattaa (kannatti vuoteen 1998 asti) siis ominaisuusdualismia. Tämä kysymys ei kuitenkaan ole tämän työn kannalta relevantti, sillä mielestäni tietoargumentista ei seuraa, että fysikalismi ontologisena oppina kumoutuisi.

mahdoton, jos hyväksyy: (1) että kaikki on fysikaalista ja (2) että kaikkea ei kuitenkaan voi selittää fysikaalisin termein.

Voidaan siis todeta, että Maryn sekundaarisella tavalla oppima uusi fakta johtaa reduktiivisen fysikalismien kumoutumiseen, mutta ei välttämättä non-reduktiivisen fysikalismien kumoutumiseen, sillä non-reduktiivinen fysikalismi kannalta ei ole mahdollista, että Mary tietäisi subjektiiviset (fysikaaliset) faktat kokematta niitä.⁵⁹

4.2. Johtopäätökset kahdeksasta kysymyksestä

Tietoargumentti on suunnattu reduktiivista a priori -fysikalismia vastaan. Tilanne kuitenkin muuttuu, jos hylkää reduktionismin, ja myöntää, ettei kaikkea fysikalismien oppiin kuuluvaa voi oppia mustavalkoisista kirjoista. Todellisuuteen kuuluu subjektiivisia faktoja, joita ei voi selittää kolmannen persoonan näkökulmasta. Tietoargumentin ensimmäinen premissi – *Marylla on kaikki fysikaalinen tieto värien havaitsemisesta ennen vapautumistaan* – ei siis ole mahdollinen. Marylla ei siis voinut olla kaikkea fysikaalista tietoa värien havaitsemisesta ennen kuin hän oli saanut ensimmäisen värikokemuksensa. Vastaus ensimmäisessä kysymykseen – *oliko Marylla todella kaikki fysikaalinen tieto ennen vapautumistaan* – on siis negatiivinen. Jäljelle jäänyt fysikalismi ei kuitenkaan ole Jacksonin esittämää reduktiivista fysikalismia, vaan jonkinlaista non-reduktiivista fysikalismia, johon kuuluvat subjektiiviset – mutta pohjimmiltaan fysikaaliset – faktat.

Voimmekin siis todeta, että jos käsittää fysikalismien non-reduktiivisesti, niin tietoargumentti kumoutuu jo ensimmäisen kysymyksen kohdalla. Mielestäni tämä onkin osuvin kritiikki tietoargumenttia vastaan. Jos kuitenkin käsittää fysikalismien reduktiivisessa mielessä, niin kritiikki ei tietenkään toimi, sillä reduktionistien mukaan fysikaaliset faktat ovat opittavissa ilman kokemuksiakin, koska kaikki on selitettävissä fysikaalisin termein. Tarkastelen siis vielä muitakin kysymyksiä, jotta reduktiivisen fysikalismien kumoutuminen tulisi selväksi.

⁵⁹ Tarkoitan non-reduktiivisella fysikalismilla tässä kantaa, jonka mukaan kaikki on (ontologisesti) fysikaalista, mutta ei välttämättä selitettävissä fysiikan avulla; toisin sanoen tietoisuutta ei voi redusoida fysikaaliseen (vaikka se onkin fysikaalinen).

Toiseen kysymykseen – *oppiiko Mary jotakin uutta* – kieltävästi vastaavat siis väittävät, että Mary ei opi vapautuessaan mitään uutta. Jos Marylla on kaikki fysikaalinen tieto, niin hän voi siitä päätellä millaista minkäkin värin havaitseminen on. Heidän mukaansa ensimmäisessä premississä sen sijaan ei ole ongelmaa – Marylla olisi kyllä (ainakin teoriassa) voinut olla kaikki fysikaalinen tieto väreistä mustavalkoisessa huoneessakin. Uuden oppimisen kieltäminen tuntuu hyvin ristiriitaiselta kannalta. Jos kuitenkin kannattaa eliminativistista oppia fenomenaalisen tietoisuuden suhteen, niin kanta on sisäisesti koherentti. Kannan hyväksyminen vaatii kuitenkin jo etukäteen päätöstä kieltää fenomenaalinen tietoisuus. (Voi kysyä mistä eliminativistit saavat oikeutukset tälle päätökselleen – mikään fysiikan laki ei tietääkseni määrää, että kaikki olisi selitettävissä fysikaalisin termein.) Tuntuukin siltä, kuin tietoargumentin puolustajat ja jyrkimmät vastustajat puhuisivat melkein eri kieltä.

If someone insists that explaining access and reportability explains everything, that Mary discovers nothing about the world when she first has a red experience [...] then I can only conclude that when it comes to experience we are on different planes. Perhaps our inner lives differ dramatically. Perhaps one of us is “cognitively closed” to the insights of the other. More likely, one of us is confused or is in a grip of a dogma. Rather, we have reached a brute clash of intuitions of a sort that is common in the discussion of deep philosophical questions. (Chalmers 1996, 167.)

Onkin siis perusteltua sanoa, että Mary oppii jotakin vapautuessaan huoneesta. Mielestäni kuitenkin seuraavat kysymykset (3, 4 ja 5) tekevät harharetken erottaessaan toisistaan faktuaalisen (*knowledge-that*) ja ei-faktuaalisen (*knowledge-how*) tiedon. Ei ole selvää, jakaantuuko tieto ylipäätään kahdenlaiseen lajiin. Tietoargumentin suhteen jako ei ainakaan ole relevantti. Sen pohjalta tehdyt päätelmät eivät siis ole päteviä. Vaikka oletettaisiinkin, että erottelu olisi mahdollista tehdä, niin taitotieto-hypoteesi ja tuttuustieto-hypoteesi kumoutuisivat joka tapauksessa: vastaesimerkit (kuten hajamielinen Mary ja sokeanäkö) osoittavat, että Mary saa uutta tietoa, joka on enemmän kuin pelkkä kyky tai tutuksi tuleminen.

Kuudennessa kysymyksessä – *saako Mary tietoonsa uusia faktoja vai vain vanhoja uudella tavalla ymmärrettynä* – on kyse faktuaalisesta tiedosta (tarkalleen ottaen edellisissä kysymyksissä muotoiltu erottelu ei ole enää merkityksellinen, kunhan ei-faktuaalinen tieto on rajattu pois). Tämä kysymys perustuu taas fregeläiseen erotteluun ilmaisun mielestä ja merkityksestä. Maryn oppima uusi fakta voisi olla samanlainen kuin tapauksessa, jossa joku saa tietää, että ”Tullius oli roomalainen”, kun hän jo tiesi, että ”Cicero oli roomalainen”. Toimiiko Maryn tapaus siis analogisella tavalla? Kysymys ei ole niin yksinkertainen kuin eräät tätä kritiikkiä kannattaneet filosofit ovat otaksuneet. Ongelma on siinä, että fenomenaalinen tieto (eli kokemus) on hyvin erilainen asia kuin kritiikeissä käytettyjen esimerkkilauseiden ilmaisema tieto. Lauseen ”Tullius oli roomalainen” olisi voinut saada selville *samalla tavalla* kuin lauseen ”Cicero oli roomalainen” (esim. historian oppikirjasta). ”Värien havaitsemisen neurologian täydellinen kuvaus mustavalkoisista lähteistä” ja ”punaisen aistikokemus” sen sijaan vaativat jollain vielä paljon jyrkemmällä tavalla erilaista esittämisen tapaa. Analogiaan perustuva argumentointi ei siis tässä tapauksessa välttämättä päde. Tästä ei kuitenkaan seuraa, etteikö Maryn saama tieto todella voisi olla tällainen vanha fakta uudeksi naamioituneena (fregeläisessä mielessä). Tämä kritiikkimuoto ei siis osoita aukottomasti tietoargumenttia vääräksi, mutta se ei kuitenkaan poissulje itsensä mahdollisuutta – tuo mahdollisuus vain tulisi osoittaa jollain toisella tavalla.

Seitsemänteen kysymykseen – *kumpi erottelutapa on oikea* – tulisi siis äskeisen perusteella kenties vastata, että Maryn tapauksessa oikea erottelutapa on *sekundaarinen tapa*. Fysikalismi ei kuitenkaan välttämättä kumoudu tähänkään, sillä on mahdollista, ettei sekundaarisellakaan tavalla ymmärretty fakta osoita, että fysikalismi olisi väärässä.

Kahdeksannessa kysymyksessä – *kumoaako Maryn sekundaarisella tavalla oppima uusi fakta fysikalismiin* – oletetaan siis, että esittämisen tapa on sekundaarinen (koska primaarisella tavalla ymmärrettynä tietoargumentti olisi jo kaatunut, eikä lisäkysymyksiä olisi tarvittu). Vaikuttaisi siltä, että Maryn oppima aidosti uusi fakta osoittaisi, että koska fysikaalinen tieto ei ollut kaikki tieto, niin fysikalismi kumoutuisi. Mutta tarkemmassa tarkastelussa selviää, että (samoista syistä kuin ensimmäisessä kysymyksessäkin) edes Maryn sekundaarisella tavalla oppima uusi

fakta ei välttämättä kumoa fysikalismia. Sillä koska Mary ei olisi yksinkertaisesti voinut tietää kaikkea fysikaalista värien havaitsemisesta ennen ensimmäistä värikokemustaan, on selvää, että Mary oppii aidosti uuden faktan saadessaan ensimmäisen värikokemuksen. Uusi fakta ei siis kumoa fysikalismia, vaan on tällaisessa tapauksessa täysin oletettavissa.

Vaikka tietoargumentti siis eräässä mielessä kumoutuisikin (koska ensimmäinen premissi ei ole mahdollinen), niin se ei silti estä tietoargumenttia kumoamasta reduktiivista fysikalismia. Reduktiivinen fysikalismihan nimittäin juuri vaatii ensimmäisen premissin (*kaiken fysikaalisen voi oppia mustavalkoisista lähteistä*) hyväksymistä. Tämäkin argumentti siis osoittaa, että fenomenaalisten faktojen ja fysikaalisten faktojen välissä on episteeminen kuilu. Sen sijaan argumentista ei seuraa, että kuilu olisi ontologinen. (Se, että argumentti *ei osoita* tätä, ei kuitenkaan riitä perusteeksi, että sitä ei olisi. Ontologinen kysymys on kuitenkin tietoargumentin ulottumattomissa.) Tietoargumentti osoittaa, että objektiivisen, kolmannen persoonan tiedon avulla ei voi selittää kaikkea. Todellisuuteen kuuluu siis subjektiivisia faktoja, jotka voi saavuttaa vain ensimmäisen persoonan näkökulmasta.

5. Fysikalistin puolustuspuheenvuoro

Tässä työssä esittämäni argumentit fysikalismia vastaan osoittavat, että fysikalismin ja tietoisuuden välissä on (ainakin) syvä episteeminen kuilu. Tarkastelen tässä lyhyesti, kuinka fysikalistit voivat vastata kolmen argumentin esittämiin syytöksiin. Fysikalisti voi esimerkiksi yrittää puolustautua kieltämällä kuilun kokonaan, tai vedota siihen, että se on pelkästään episteeminen, jolloin ainakin ontologinen fysikalismi säilyttäisi asemansa. Chalmersin mukaan (2002a, 251–260) fysikalistilla on karkeasti ottaen kolme tapaa suhtautua fysikaalisen ja fenomenaalisen väliseen kuiluun: hän kutsuu niitä A- B- ja C-tyypin materialismeiksi.⁶⁰

⁶⁰ Käytän Chalmersin luomaa erottelua tässä jokseenkin vapaasti ja suuntaa antavasti. Chalmers itse ei tietenkään ole minkäänlaisen materialismin kannalla, vaan hän muotoilee lisäksi myös kolme ei-materialistista kantaa: D-tyypin dualismin (interaktionismi), E-tyypin dualismin (epifenomenalismi) ja F-tyypin monismin (panprotopsykismi).

A-tyypin materialistien mukaan fysikaalisten ja fenomenaalisten totuuksien välissä ei ole minkäänlaista (episteemistäkään) kuilua. Tällaisessa materialismissa *ei ole* tietoisuuden vaikeaa ongelmaa. Kun on selvittänyt helpot ongelmat – kognitiivisuuteen, käyttäytymiseen, ympäristöön jne. liittyvät psykologisen pääsytietoisuuden faktat – on tietoisuus selvitetty kokonaan. Tämän kannan mukaan ei ole (loogisestikaan) mahdollista kuvitella fysikaalisia kaksoisolentoja, joilla ei olisi lainkaan tietoisuutta. Jos on Daniel C. Dennettin fysikaalinen kaksoisolento, on Daniel C. Dennettin kaksoisolento joka suhteessa. Mary taas ei opi mitään uutta mustavalkoisesta huoneesta vapautuessaan (eli A-materialisti vastaa kieltävästi tietoargumentin kritiikkien toiseen kysymykseen). A-materialistin mukaan Sokeiden Valtakunnan vuorikiipeilijän olisi täysin mahdollista saada sokeat ymmärtämään, mitä on näkeminen, kunhan hänellä olisi vain tarpeeksi aikaa ja tieteellistä tietoa (ja sokeilla tarpeeksi ymmärryskykyä). Jos hän selittäisi heille juurta jaksain, kuinka silmän hermosolut reagoivat valoon ja välittävät informaation näköaivokuorelle ja muualle aivoihin, hän olisi selittänyt heille *kaiken* näkemisestä.⁶¹

A-materialisti voi olla eliminativisti, ja kieltää kokonaan tietoisuuden sekä fenomenaalisten ominaisuuksien olemassaolon. Toinen mahdollisuus on myöntää tietoisuuden olemassaolo käyttämällä tietoisuuden käsitettä siten, että se on täysin selitettävissä funktionalistisilla ja behavioristisilla termeillä. Tällaisessa tapauksessa tietoisuus voisi olla esimerkiksi vain sitä, että on pääsy johonkin informaatioon tai että on kykeneväinen verbaaliseen kanssakäymiseen. A-tyypin materialistille tietoisuus on siis vain psykologista pääsytietoisuutta. Tässä työssä esitetyt fysikalismin kritiikit, jotka perustuvat fenomenaalisen tietoisuuden poikkeukselliseen luonteeseen, eivät kosketa A-materialisteja, sillä heille fenomenaalista tietoisuutta ei ole lainkaan olemassa. (Tai tällä tavoin he haluavat ajatella.)

Daniel Dennettin (1991; 2003) *heterofenomenologia* on yksi esimerkiksi A-tyypin materialismista: siinä verbaalinen raportointi (”kertomukset”) riittää tietoisuuden kriteeriksi. Kvalioita ei tarvita; tarvitsee vain tutkia raportteja, joissa subjekti kertoo omasta mentaalisestä tilastaan. Raportin tutkija pääsee tällä tavoin kolmannen persoonan näkökulmasta käsiksi raportissa esitettyyn ensimmäisen persoonan

⁶¹ Sokeiden Valtakunnan esimerkki toimii tässä siis lähes samalla tavalla kuin tietoargumentti.

näkökulmaan. Tutkija ei vain kuuntele, vaan hän lisäksi tarkkailee muutenkin tutkittavana olevaa subjektia: tämän ulkoista käytöstä, eleitä, ilmeitä jne. Kyseessä on siis hyvin paljon behaviorismia muistuttava kanta. Heterofenomenologian keinoin zombi-kysymys on merkityksetön: haastatteleamalla olioita ei voida tehdä mitään periaatteellista eroa potentiaalisten zombien ja ei-zombien välille.

The total set of details of heterophenomenology, plus all the data we can gather about concurrent events in the brains of subjects and in the surrounding environment, comprise the total data set for a theory of human consciousness. It leaves out no objective phenomena and no subjective phenomena of consciousness. (Dennett 2003, 2.)

A-tyyppin materialismin ratkaisu tietoisuuden ongelmaan on siis kieltää ongelman olemassaolo. Mutta vaikka A-materialisti selittäisi tyhjentävästi kaikki psykologisen pääsy tietoisuuden ongelmat (eli hänen mukaansa *kaikki* tietoisuuden ongelmat), on jotain mikä jää täysin käsittelemättä – miksi kokemus tuntuu jonkinlaiselta. Kanta on mielestäni siis epäonnistunut: se joko kieltää kysymyksen, tai vastaa aivan eri kysymykseen kuin mikä oli asetettu.

B-tyyppin materialistien mukaan fysikaalisten ja fenomenaalisten faktojen välissä on episteeminen kuilu, mutta ei ontologista kuilua. Tämän kannan mukaan puuttuvat ja käänteiset kvaliat ovat kuviteltavissa, mutta se ei tee niistä metafysisesti mahdollisia. Daniel C. Dennettin fysikaalinen kaksoisolento ilman tietoisuutta on kuviteltavissa, mutta ei mahdollinen. Mary oppii kyllä jotain, mutta kyseessä ei ole fakta, vaan vain uusi taito, tuttuus tai vanha fakta vain uudella tavalla esitettynä. Vuorikiipeilijä ei voi selittää sokeille tyhjentävästi, millaista on näkeminen, mutta tuo tieto itsessään kuitenkin koskee pelkästään (ontologista) fysikaalista maailmaa. Jos joku sokeista saisi yllättäen näkökyvyn, hän oppisi jotain siinä mielessä, että hän saisi uuden kyvyn tai taidon. Sinänsä uutta tietoa hän ei saisi, jos hän jo tietäisi kaiken fysikaalisen. B-materialistit hyväksyvät, että tietoisuuden helppojen ongelmien lisäksi on olemassa myös tietoisuuden vaikea ongelma, mutta se ei tarkoita, että helppojen ja vaikeiden ongelmien välillä olisi ontologinen ero. Tällaisen materialismin mukaan tietoisuutta ei voi siis selittää (episteemisesti) fysikaalisilla ja funktionaalisilla termeillä, vaikka (ontologisesti) kyseessä on aivan samanlainen

fysikaalinen ilmiö kuin kaikki muukin. Esimerkiksi Levine (1983) edustaa selityksellisen kuilun argumentillaan juuri tällaista kantaa.

Chalmers (2002a, 253) luokittelee myös identiteettiteorian B-materialismiksi: tietoisuuden käsite on kyllä erilainen (episteemisesti) kuin fysikaaliset käsitteet, mutta tarkempi tutkimus osoittaa, että käsitteet ovat identtisiä (ontologisesti), siis viittaavat samaan asiaan (vain *eri tavalla*). Tämä mahdollistaa kannan, joka ottaa huomioon tietoisuuden käsitteen episteemiset ongelmat, mutta säilyttää fysikalistisen maailmankuvan. Kanta ei vain ota huomioon sitä, että tietoisuutta ei voi redusoida samalla tavalla kuin muita maailman ilmiöitä. Chalmers (2002a, 254) huomauttaa myös, että B-materialisti voisi vedota Kripken teoriaan välttämättömästi a posteriorista. Kenties tietoisuuden ja fysikaalisen välinen suhde onkin samanlainen aposteriorinen välttämättömyys kuin veden ja H₂O:n välinen suhde. Silloin olisi siis syytäkin olettaa, että ilmiöiden välissä on episteeminen kuilu (koska suhde ei ole a priori), mutta ontologisen kuilun vaaraa ei ole (koska suhde on välttämätön).⁶² Chalmersin mukaan tietoisuus ei tässäkään tapauksessa toimi samalla tavalla kuin muut ilmiöt: veden ja H₂O:n välinen suhde voi olla aposteriorinen, sillä sen voi johtaa maailman fysikaalisesta kuvauksesta. Tietoisuuden ja fysikaalisen välistä suhdetta ei voi kuitenkaan johtaa (aposteriorisesti) fysikaalisesta kuvauksesta, sillä kuvauksessa ei ole mitään, mikä sen selittäisi. Episteeminen kuilu on välissä.

Chalmers on yhdistänyt B-materialismiksi niin monia erilaisia fysikalistisia kantoja, että siitä on vaikea sanoa, onnistuuko se kokonaisuudessaan puolustamaan fysikalistista oppia. Yleisesti voisi kuitenkin todeta, että tietoisuuden ja fysikaalisen välisen episteemisen kuilun samaistaminen muiden alojen episteemisiin kuiluihin on epäonnistunut yritys. Tietoisuuden ja fysikaalisen välinen kuilu ei ratkea analogisesti muiden kuilujen kanssa (kuten esim. Kripke on osoittanut). B-materialistien yrittämät analogia-ratkaisut eivät toimi sen takia, että niissä ei oteta huomioon tietoisuuden (fenomenaalista) luonnetta, joten niiden kuvaama tietoisuus ei enää olekaan sama tietoisuus, joka alun perin oli selitettävänä. Tällä tavalla tullaankin selittäneeksi vain

⁶² Van Gulickin (2003) mukaan kuilu on odotettavissa, ja se kuuluu ontologiseen fysikalismiin. ”The pragmatic contexts in which our subjective concepts are embedded are so unlike those from which our third-person concepts derive their contents that a gap between the two is very likely, indeed all but inevitable (Van Gulick 2003, 346).

psykologisen pääsytietoisuuden kysymyksiä. Siispä B-materialismikaan ei selviä esittämistäni fysikalismia vastustavista argumenteista.

C-tyypin materialistit hyväksyvät, että fysikaalisen ja fenomenaalisen välissä on syvä episteeminen kuilu, mutta heidän mukaansa se on ainakin periaatteessa ylitettävissä. Tämän kannan mukaan Daniel C. Dennettin zombi-kaksoisolento on kuviteltavissamme *tällä hetkellä*, mutta ei loputtomasti – ei enää siinä vaiheessa kun olemme oppineet tarvittavan määrän lisää tietoisuudesta – kun kuilu on ylitetty. Puuttuva ja käänteinen kvalia ovat Chalmersin termeillä (2002a; 2002b) *prima facie* -kuviteltavissa, mutta eivät *ideaalisesti* kuviteltavissa. Vaikka nyt vaikuttaa mahdottomalta ratkaista tietoisuuden vaikea ongelma fysikaalisilla termeillä, se on kuitenkin *ratkaistavissa*. Marylla ei voi mustavalkoisessa huoneessaan *vielä* olla kaikkea fysikaalista tietoa värien havaitsemisesta (eli C-tyypin materialisti vastaa kieltävästi kritiikkien ensimmäiseen kysymykseen), mutta kunhan tiede kehittyy, hän voi mustavalkoisessa huoneessaankin tietää, millaiselta punainen näyttää. Vuorikiipeilijä ei voi selittää sokeille, millaista on näkeminen, koska hänellä ei nykytieteen rajoittuneisuuden takia ole siihen keinoja. Kun tiede (tässä tapauksessa näköhavainnon fysiologia) vain on tarpeeksi kehittynyttä, hän voi oikein hyvin saada sokeat ymmärtämään, mitä on näkeminen pelkästään kertomalla heille, miten näköjärjestelmä toimii.

Chalmers myöntää, että kanta on houkutteleva (vaikkakin hänen mukaansa väärässä). Siitä on hieman erilaisia versioita. Nagel (1974) esimerkiksi on esittänyt, että samalla tavalla kuin esisokraatikot eivät voineet ymmärtää, kuinka aine voi olla energiaa, emme mekään nykyisen tietomme avulla voi ymmärtää, kuinka tietoisuus voi olla fysikaalista. Tietämättömyys ei kuitenkaan estä ainetta olemasta energiaa – eikä tietoisuutta fysikaalista. McGinnin (1989) mukaan taas voi olla mahdollista, ettei ihmisymmärrys koskaan kykenekään ratkaisemaan ongelmaa, mutta se johtuu vain ymmärryksemme rajoituksista – ongelma sinänsä on kyllä periaatteessa ratkaistavissa.

C-materialistit saattavat vedota siihen, että kuilu johtuu vain siitä, että fysiikka ei ole vielä tarpeeksi kehittynyttä. Nykyinen rajoittunut käsityksemme maailmasta (eli fysiikasta) saa tietoisuuden ja fysikaalisen välisen kuilun vaikuttamaan

ylitsepääsemättömältä, mutta todellisuudessa, ”ulkomaailmassa” ei vallitse mitään kuilua. Chalmers (2002a, 258) kuitenkin tyrmää ajatuksen: fysiikka (myös ideaalinen, aukoton fysiikka) käsittelee rakenteita ja dynamiikkaa, eikä niistä ole selittämään fenomenaalisen ja fysikaalisen välistä kuilua, koska kuilussa ei ole kyse rakenteista ja dynamiikasta. Joku voisi väittää, että voisi olla ”uudenlaisia” fysikaalisia teorioita, jotka ulottuvat rakenteiden ja dynamiikan ”ulkopuolelle”. Mutta fysikaalisen selittämisen *luonteeseen* kuuluu, että uudet fysikaaliset ominaisuudet ja ilmiöt postuloidaan aina suhteessa niiden kykyyn selittää rakenteita ja dynamiikkaa. Muunlainen selitys ei olisi enää fysiikkaa.

C-materialisti voi kuitenkin lähestyä ongelmaa toiseltakin puolelta: kenties ei tarvitsekaan kasvattaa fysiikan tietämystä postuloimalla uusia fysikaalisia ominaisuuksia, jos nykyinenkin fysiikka voi selittää jo suoraan tietoisuuden. Tällainen ratkaisu on esimerkiksi kvanttimekaniikkaan vetoaminen: perinteisellä fysiikalla ei voi selittää tietoisuutta, mutta kvanttifysiikalla voi. Kuilu on siis jo ylitetty, ja enää tarvitsee vain tarkentaa kvanttifysiikan tuntemusta, joka on kattava selitys myös tietoisuudelle. Chalmersin (2002a, 259) mukaan kuitenkin tällainenkin kanta johtaa materialistin kannalta ongelmalliseen lopputulokseen. Jos hyväksyy, että suoraan tietoisuuden olemukseen vetoava fysiikan laji (kuten kvanttimekaniikka) *on tietoisuuden selitys*, ei tietoisuus redusoidu muuhun (kuten perinteiseen ei-quantti-) fysiikkaan. Seurauksena on siis jonkinlainen non-reduktiivinen kanta. Chalmersille voisi huomauttaa, että tämä ei varmaankaan ole ongelma kaikille kvanttiteorian kannattajille; he itse saattavat kyllä hyväksyä kantansa johtamisen non-reduktiiviseen selitykseen tietoisuudesta. Mutta silloin kyse ei kuitenkaan ole enää Chalmersin kuvailemasta C-tyyppin materialismista.

C-tyyppin materialismi ei ole riippumaton muista kannoista: kun (tulevaisuuden hetkenä x) saamme selville mikä tietoisuuden ja fysikaalisen suhde todella on, täytyy C-tyyppistä luopua. Kun kuilu on ylitetty, on tietämättömyys loppunut, ja selitys (oli se sitten minkälainen tahansa) tullut tilalle. Oletetun selviämishetken jälkeen C-tyyppin materialistin täytyy siis liittyä A- tai B-tyyppin materialisteihin (tai dualisteihin).

Väittäisin siis, että C-materialismistakaan ei siis ole fysikalismin puolustajaksi. Tämä johtuu siitä, että se ei loppujen lopuksi ole itsenäinen kanta, vaan osoittautuu pohjimmiltaan A- tai B-tyyppin materialismiksi, jotka puolestaan jo osoittautuivat huonoiksi fysikalismin puolustajiksi. Onko fysikalismi siis väärässä? Tietoisuuden *selittämisen* suhteen: kyllä on.

Fysikalismia ylipäätään vaivaa ns. *Hempelin dilemma* (Stoljar 2009), jonka mukaan fysikalismi on joko väärässä tai triviaalia. Jos fysikalismi määritellään suhteessa nykyfysiikkaan, se on *väärässä*, sillä eihän kukaan väitäkään että nykyfysiikka olisi (joka suhteessa) oikeassa. Jos se taas määritellään suhteessa tulevaisuuden fysiikkaan, niin se on *triviaalia*, sillä kukaan ei voi tietää, mitä se sisältää. Fysikalismi on myös eräänlainen kehäpäätelmä: kaikki ilmiöt ovat selitettävissä fysiikan avulla, koska fysiikka (oikein muotoiltuna) on se asia, *joka selittää kaikki ilmiöt*. Jos siis väitetään, että tietoisuus on fysikaalinen, niin fysikalismi selittää sen ilman muuta, sillä ”täydellinen fysiikka” selittää kaiken. Selitys on kuitenkin triviaali, ja fysikalismi vaatisi jonkin kehän ulkopuolelta tulevan perustelun.

6. Lopuksi

Kaikkien kolmen tässä työssä esitetyn argumentin perusteella voidaan siis todeta, että fysikalismi ei ole oikea tapa selittää tietoisuutta. Fysikalismin ei tarvitse kuitenkaan olla ontologisesti väärä kanta, sillä episteeminen puute ei lopulta paljasta muuta kuin *episteemisen puutteen*. Vaikka tietoisuutta ei kyettäisi koskaan selittämään fysikaalisilla termeillä, ei fysikalistisen ontologian tarvitse vaarantua. On kapeakatseista olettaa, että juuri fysiikan pitäisi kyetä selittämään *kaikki*. Oletus ei perustu mihinkään universaaliin lakiin. Mutta vaikka reduktiivinen fysikalistinen selitys epäonnistuukin, voi tietoisuus silti olla pohjimmiltaan fysikaalinen. Jos hyväkyy fysikalistisen ontologian, mutta ei reduktiivista selittämistä, tulee siis kannattaneeksi jonkinlaista non-reduktiivista fysikalismia tai ominaisuusdualismia.

Argumentit osoittavat todella vakuuttavasti, että tietoisuuteen kuuluu faktoja, jotka eivät ole selitettävissä fysikaalisin termein. Todellisuuteen kuuluu siis subjektiivisia

kokemuksellisia faktoja, joista olemme tietoisia siinä mielessä, että niiden kokeminen tuntuu jonkinlaiselta. Niiden luonteeseen kuuluu, että niistä voi saada tietoa vain ensimmäisen persoonan näkökulmasta. Jos nämäkin faktat kuitenkin ovat siis pohjimmiltaan fysikaalisia (koska episteemisen argumentin hyväksyminen ei vaadi ontologisesta fysikalismista luopumista), millä tavalla niistä voi saada tietoa? Millä tavalla voi saada tietoa tietoisuuden subjektiivisista fysikaalisista ominaisuuksista, jotka ovat ymmärrettävissä vain jostain tietystä näkökulmasta? Tähän kysymykseen vastaamisen estää selityksellinen kuilu.

What creates the philosophical puzzlement is the assumption that the problem must somehow be scientific but that any science *we* can come up with will represent things as utterly miraculous. And the solution is to recognize that the sense of miracle comes from us and not from the world. There is, in reality, nothing mysterious about how the brain generates consciousness. There is no *metaphysical* problem. (McGinn 1989, 401.)

Selityksellinen kuilu tekee tietoisuudesta poikkeuksellisen, *ihmeellisen*, ilmiön kaikkiin muihin maailman ilmiöihin verrattuna. Joskus onkin vaikea muistaa, että tietoisuuden mysteeri ei ole metafysiikan ongelma, vaan vain sen ymmärryksemme ongelma.

Entä mitä tapahtui Sokeiden Valtakunnassa? Bogota tottui lopulta yhteisön tapoihin, ja rupesi itsekin epäilemään, onko hän vain kuvitellut koko ulkomaailman. Hän rakastui sokeaan tyttöön nimeltään Medina-saroté, mutta ei saanut naida tätä, koska oli yhteisön silmissä hullu. Yksi kylän vanhimmista kuitenkin oivalsi, mikä Bogotaa vaivaa: hänellä oli kasvoissaan silmiksi kutsutun alueen kohdalla omituiset kasvaimet, jotka häiritsivät hänen aivotoimintaansa ja aiheuttivat mielisairauden. Yksinkertaisella leikkauksella, jossa poistettaisiin nuo Bogotan sairaalloiset osat, saataisiin hänestä tervejärkinen ja täysivaltainen kansalainen. Bogota suostui leikkaukseen, sillä se oli hänen ainoa keinonsa saada Medina-saroté omakseen. Viime hetkellä hän kuitenkin pakeni takaisin vuorille nähdessään auringonnousun kaikessa väriloistossaan – jonka voi ymmärtää vain subjektiivisen kokemuksen avulla.

Lähteet:

Alter, Torin (2001): ”Know-How, Ability, and the Ability Hypothesis”. *Theoria* 67, s. 229–239.

Bickle, John (2008): ”Multiple Realizability”. *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (toim.), URL: <http://plato.stanford.edu/archives/fall2008/entries/multiple-realizability/>.

Bigelow, John & **Pargetter**, Robert (1990): ”Acquaintance with Qualia”, Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There’s Something About Mary*. Cambridge: MIT Press, 2004, s. 179–195.

Block, Ned (2002): ”Concepts of Consciousness”. Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*. New York: Oxford University Press, 2002, s. 206–218.

Campbell, Keith (1970): *Body and Mind*. London: Macmillan.

Chalmers, David J. (1995): ”Facing Up to the Problem of Consciousness”. Saatavilla www.muodossa.net: URL: <http://consc.net/papers/facing.html/>. Luettu 16.04.2012.

Chalmers, David J. (1996): *The Conscious Mind, In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, David J. (2002a): ”Consciousness and Its Place in Nature”. Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*, New York: Oxford University Press, 2002, s. 247–272.

Chalmers, David J. (2002b): ”Does Conceivability Entail Possibility?”. Teoksessa Gendler & Hawthorne (toim.): *Conceivability and Possibility*, Oxford: Oxford University Press, 2002, s. 145–200.

Churchland, Paul M. (1985): ”Reduction, Qualia, and the Direct Introspection of Brain States”. *Journal of Philosophy* 82, s. 8–28.

Conee, Earl (1994): ”Phenomenal Knowledge”, Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There’s Something About Mary*. Cambridge: MIT Press, 2004, s. 197–215.

Davidson, Donald (1970): ”Mental Events”. Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*. New York: Oxford University Press, 2002, s. 116–125.

Dennett, Daniel C. (1991): *Consciousness Explained*. Boston: Little Brown and Company.

- Dennett**, Daniel C. (2003): "Who's On First? Heterophenomenology Explained". *Journal of Consciousness Studies* 10, No. 9–10, s. 19–30.
- Descartes**, René (2001): *Metodin esitys*. Teoksessa Jansson (suom.): *René Descartes, Teokset I*, Helsinki: Gaudeamus, 2001, s. 117–168. Alkuperäisteos *Discours de la Méthode* (1637).
- Fodor**, Jerry A. (1974): "Special Sciences (Or the Disunity of Science as a Working Hypothesis)". *Synthese* 28, s. 97–115.
- Funkhouser**, Eric (2007): "Multiple Realizability". *Philosophy Compass* 2/2, s. 303–315.
- Gendler**, Tamar Szabó & **Hawthorne**, John (2002): "Introduction: Conceivability and Possibility". Teoksessa Gendler & Hawthorne (toim.): *Conceivability and Possibility*, Oxford: Oxford University Press, 2002, s. 1–70.
- Jackson**, Frank (1982): "Epiphenomenal Qualia". *Philosophical Quarterly* 32, s. 127–136.
- Jackson**, Frank (1986): "What Mary Didn't Know". *Journal of Philosophy* 83, s. 291–295.
- Jackson**, Frank (2002): "Mind and Illusion". Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*, 2004, Cambridge: MIT Press, s. 421–442.
- Kalat**, James W. (2004): *Biological Psychology*. Belmont: Wadsworth.
- Kim**, Jaegwon (1992): "Multiple Realization and the Metaphysics of Reduction". Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*. New York: Oxford University Press, 2002, s. 135–149.
- Kim**, Jaegwon (2006): *Philosophy of Mind*. Cambridge: Westview Press.
- Kirk**, Robert (1974): "Sentience and Behaviour". *Mind, New Series* 83, No. 329, s. 43–60.
- Kripke**, Saul (1980): *Naming and Necessity*. Cambridge: Harvard University Press.
- Levine**, Joseph (1983): "Materialism and Qualia: The Explanatory Gap". *Pacific Philosophical Quarterly* 64, s. 354–361.
- Lewis**, David (1988): "What Experience Teaches". Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*. Cambridge: MIT Press, 2004, s. 77–103.
- Loar**, Brian (1997): "Phenomenal States (Revised Version)", Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*, 2004, Cambridge: MIT Press, s. 219–239.

Locke, John (1975): *An Essay Concerning Human Understanding*. Oxford: Oxford University Press. Alkuperäisteos vuodelta 1689.

McGinn, Colin (1989): "Can We Solve the Mind-Body Problem?". Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*. New York: Oxford University Press, 2002, s. 394–405.

Montero, Barbara (1999): "Frank Jackson Changed His Mind". Philosophy News Service.

Nagel, Thomas (1974): "What is it like to be a bat?". Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*. New York: Oxford University Press, 2002, s. 135–149.

Nemirow, Lawrence (1990): "Physicalism and the Cognitive Role of Acquaintance". Teoksessa Lycan (toim.): *Mind and Cognition: A Reader*. Oxford: Blackwell, 1990, s. 490–499.

Nida-Rümelin, Martine (1996): "Pseudonormal Vision: An Actual Case of Qualia Inversion?". Teoksessa Chalmers (toim.): *Philosophy of Mind, Classical and Contemporary Readings*. New York: Oxford University Press, 2002, s. 99–105

Putnam, Hilary (1967): "The Nature of Mental States". Teoksessa Putnam: *Mind, Language and Reality: Philosophical Papers, Vol 2*. New York: Cambridge University Press, 1975, s. 429–440.

Robinson, Howard (1993): "Dennett on the Knowledge Argument". Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*. Cambridge: MIT Press, 2004, s. 69–73.

Ryle, Gilbert (1949): *The Concept of Mind*. Harmondsworth: Penguin Books Ltd, 1973.

Stoljar, Daniel & Nagasawa Yujin (2004): "Introduction". Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*. Cambridge: MIT Press, 2004, s. 1–36.

Stoljar, Daniel (2009): "Physicalism". The Stanford Encyclopedia of Philosophy (Fall 2009 Edition), Edward N. Zalta (toim.), URL: <http://plato.stanford.edu/archives/fall2009/entries/physicalism/>.

Shoemaker, Sydney (1982): "The Inverted Spectrum". *Journal of Philosophy* 79, s. 357–381.

Tye, Michael (2000): "Knowing What It Is Like: The Ability Hypothesis and the Knowledge Argument". Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*. Cambridge: MIT Press, 2004, s. 143–160.

Van Gulick, Robert (1993): "Understanding the Phenomenal Mind: Are We All just Armadillos?". Teoksessa Davies & Humphreys (toim.): *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell, 1993, s. 137–154.

Van Gulick, Robert (2003): "Maps, Gaps and Traps". Teoksessa Smith & Jokic (toim.): *Consciousness, New Philosophical Perspectives*. Oxford: Oxford University Press, 2003, s. 323–352.

Van Gulick, Robert (2004): "So Many Ways of Saying No to Mary". Teoksessa Ludlow, Nagasawa & Stoljar (toim.): *There's Something About Mary*. Cambridge: MIT Press, 2004, s. 365–405.

Wells, H. G. (1904): "The Country of the Blind", teoksessa *The Country of the Blind and Other Stories*. Leipzig: Bernhard Tauchnitz, 1926, s. 7–48.