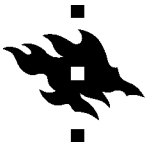


Tietoisuuden vaikea ongelma

ja intuitiot joille se perustuu

Jouko Kivinen
Pro Gradu -tutkielma
Teoreettinen filosofia
Humanistinen tiedekunta
Helsingin yliopisto
Maaliskuu 2019



Tiedekunta/Osasto – Fakultet/Sektion – Faculty Humanistinen tiedekunta		Laitos – Institution – Department Filosofian, historian, kulttuurin ja taiteiden tutkimuksen laitos	
Tekijä – Författare – Author Jouko Kivinen			
Työn nimi – Arbetets titel – Title Tietoisuuden vaikea ongelma ja intuitiot joille se perustuu			
Oppiaine – Läroämne – Subject Teoreettinen filosofia			
Työn laji – Arbetets art – Level Pro Gradu -tutkielma		Aika – Datum – Month and year Maaliskuu 2019	Sivumäärä– Sidoantal – Number of pages 88
Tiivistelmä – Referat – Abstract <p>Tutkielma käsittelee analyttisen perinteen edustajan David Chalmersin 1994 esittelemää tietoisuuden ”vaikeaa ongelmaa” ja sen perustavia intuitioita. Chalmersin mukaan tietoisuuden suhteen reduktionistiset teoriat eivät kykene selittämään, miksi tietoisuuden funktioihin ja prosesseihin liittyy lisäksi kokemuksellinen tietoisuus. Chalmersin mukaan kokemuksellisuus on intuitiivisesti kiistämätön ja episteemisesti ensisijainen. Tälle ja joillekin muille intuitioille Chalmers rakentaa realistisen tietoisuuden teoriansa, jota hän kutsuu ominaisuusdualismiksi.</p> <p>Tätä Chalmersin ainakin vielä 2000-luvun alussa kannattamaa kantaa vastaan asetan illusionis-miksi nimetyn kannan, jonka mukaan tietoisuus on illuusio. Tämä on Chalmersin luokittelussa A-tyyppin materialistinen, myös eliminativistiseksi sanottu näkemys. Suuntauksen tunnetuin edustaja on Daniel Dennett, jonka Chalmersin ja muiden realistien filosofiaan kohdistama kritiikki täyttää merkittävän osan tutkielmasta.</p> <p>Tutkielma on metodiltaan lähdekirjallisuuden vertaileva ja kriittinen tutkimus. Lähdemateriaalina käytän erityisesti Chalmersin ja Dennettin artikkeleita ja muita kirjoituksia 90-luvulta nykypäivään, lisäksi muita pääasiassa saman ajanjakson keskeisiä kirjoituksia.</p> <p>Luvussa 1 esittelen Chalmersin lähtökohdan ja keskeiset antagonistit, luvussa 2 esittelen materialistisen näkemyksen ja sen ongelmat. Luvuissa 3 ja 4 käyn läpi Chalmersin ratkaisuyrityksen. Luvussa 5 luettelin intuitiot, joiden varaan Chalmersin ratkaisuyritys rakentuu, ja luvut 6–8 käsittelevät illusionismia ja erityisesti kritiikkiä, jonka Dennett kohdistaa realistiseen näkemykseen. Luvuissa 9 ja 10 teen yhteenvedon em. kritiikistä ja esitän joitakin loppupäätelmiä.</p> <p>Totean, että Chalmersin ominaisuusdualismi perustuu kestäättömässä määrin perustelemattomille intuitioille. Esitän, että vaikea ongelma on luultavasti näennäisongelma, joka perustuu kognitiivisten prosessien introspektiiviseen läpinäkymättömyyteen, ja lisäksi että – varsinkin dualistien käyttämät – käsitteet kuten ”tietoisuus”, ”kvalia” ja ”millaista-on-olla” ovat liian epämääräisiä tieteelliseen tutkimukseen. Lopuksi spekuloin, että tietoisuus on ehkä aikanaan selitettävissä yhdistelmänä ennakoivan prosessoinnin teoriaa, introspektiivista läpinäkymättömyyttä ja mahdollisesti hermoverkkotutkimuksesta saatavia malleja.</p>			
Avainsanat – Nyckelord – Keywords Tietoisuus, kvalia, vaikea ongelma, intuitio, dualismi, mielenfilosofia, analyttinen filosofia			
Säilytyspaikka – Förvaringställe – Where deposited Keskustakampanin kirjasto			
Muita tietoja – Övriga uppgifter – Additional information			

Sisällys

1	Johdanto	3
1.1	Tutkimusongelma	3
1.2	Tutkielman rakenne	5
2	Taustaa	6
2.1	Materialismi	6
2.2	Tietoisuuden muodostama ongelma	8
3	Argumentteja materialismia vastaan	9
3.1	Kaksiulotteinen semantiikka	10
3.2	Selitysargumentti	11
3.3	Kuviteltavuusargumentti	14
3.4	Tietoargumentti	17
4	Chalmersin tietoisuusteoria	19
4.1	Metodologiset periaatteet	20
4.2	Psykofyysiset periaatteet ja lait	21
4.3	Epifenomenalismi	25
4.4	Fenomenaaliset arvostelmat	27
5	Realistien keskeiset intuitiot	32
6	Tietoisuus illuusiona	34
7	Dennett	39
7.1	Heterofenomenologia	39
7.1.1	Tutkimusongelma	39
7.1.2	Metodi	41
7.1.3	Tulkinta	43
7.2	Dennettin tietoisuusteoria	45
8	Dennettin kritiikki	47
8.1	Metafyysiset intuitiot	49
8.2	Selittävät intuitiot	52
8.3	Tietointuitiot	54
8.4	Modaali-intuitiot	57
8.4.1	Zombit ja invertoidut kvaliat	57
8.4.2	Argumentin premissien ja validiuden kritiikki	59

8.5	Itseä koskevat intuitiot.....	61
8.6	Kvaliteetti-intuitiot	63
8.7	Esitysintuitiot.....	67
8.8	Arvointuitiot	70
8.9	Kattavuusintuitiot	71
9	Yhteenveto Dennettin kritiikistä.....	71
10	Yhteisen maaperän etsintää	75
	Lähteet.....	80

1 Johdanto

1.1 Tutkimusongelma

David J. Chalmers esitti Tucsonissa 1994, *Toward a Science of Consciousness* –konferenssissa huomiota herättäneen jaottelun tietoisuuden ”helppoihin” ja ”vaikeaan” ongelmaan, jota hän kehitteli lisää kirjassaan *The Conscious Mind* (1996). Tietoisuuden helpot ongelmat koskevat esimerkiksi oppimista, muistia, introspektiota, kykyä raportoida omista mielentiloistaan, tarkkaavaisuutta jne. (Chalmers 1995, 200; 1996, 6–11). Näitä kaikkia voidaan tutkia empiirisesti *kolmannen persoonan näkökulmasta* vaikkapa kyselytutkimusten, fMRI- tai varjoainekuvausten avulla, ja ne voidaan periaatteessa kuvata funktionaalisilla tai komputatiivisilla malleilla.

Vaikea ongelma sen sijaan koskee *kokemusta*: miksi näihin funktioihin ja fysikaalisiin prosesseihin näyttää liittyvän lisäksi erityinen kokemus siitä millaista on olla minä, punaisen ”punaisuus” tai kokemus siitä, että *minuun sattuu*? Tätä tietoisuuden tyyppiä kutsutaan yleensä *fenomenaaliseksi* tai *kokemukselliseksi* tietoisuudeksi (Chalmers 1995, 201; 1996, xii) – erotuksena helppojen ongelmien käsittelemästä *psykologisesta* tietoisuudesta – ja sitä koskee erityinen *ensimmäisen persoonan* näkökulmasta saatava havaintoaineisto (Chalmers 2004a, 1111). Fenomenaaliseen mielentilat *tuntuvat* joltakin, niillä on *laadullinen* aspekti, esimerkiksi *vihreys kuten minä sen koen*, ja näitä ominaisuuksia nimitetään yleisesti *kvalioiksi*.

Sekä ensimmäisen että kolmannen persoonan havainnot tarvitsevat selityksen (Chalmers 2004a, 1111). Etsitty vaikean ongelman ratkaisu käsittäisi kuvauksen tietoisuuden ja fysikaalisten prosessien suhteesta ja selityksen siitä miten fysikaaliset prosessit liittyvät kokemuksellisiin tiloihin (Chalmers 2002a, 248).

Tutkielman päähuomio on Chalmersin 1990-luvun ja seuraavan vuosikymmenen alun ajattelussa, jossa Chalmers tukeutuu monin paikoin intuitioon – tai arkipsykologisiin käsityksiin, kuten hänen kriitikonsa sanovat. Selitettyyään funktionalistisia ja fysikaalistisia tietoisuuden teorioita Chalmers lisää aina kysymyksen *mutta miksi on lisäksi kokemusta?* – kysymys joka tuntuu käyvän Chalmersille kritiikiksi mitä tahansa tietoisuuden teoriaa vastaan. Samoin kvalioiden luonteesta hänellä on vahva, selvä intuitiivinen käsitys: ne eivät ole reduktiivisesti selitettävissä. Edelleen, Chalmersille tietoisuuden

episteeminen ensisijaisuus on jotakin jota ei voi epäillä, *explanandum*, joka pitää ottaa vakavasti (Chalmers 1995, 209; 1996, xii-xiii). Chalmersin mukaan materialismin puolustajat eivät yleensä onnistu selittämään tietoisuuden vaikeaa ongelmaa, vaan he selittävät joitakin tietoisuuden funktionaalisia ominaisuuksia tai suoraan kieltävät vaikean ongelman olemassaolon kokonaan (Chalmers 1995, 206).

Tietoisuuden vaikean ongelman kieltävää kantaa on kutsuttu myös *illusionismiksi*, erotuksena Chalmersin ja hänen kannattajiensa *realismista*. Illusionismin heikossa muodossa fenomenaalinen tietoisuus käsitetään olemassa olevaksi, mutta niin, että jotkin sitä koskevat intuitiomme ovat illuusioita (kuten käsitys tietoisuudesta primitiivisenä ja ei-fysikaalisena). Illusionismin vahvassa muodossa tietoisuuden vaikea ongelma ja tietoisuus itse käsitetään illuusioksi.

Realistien ja illusionistien näkemysten välinen ero on erittäin suuri. Chalmersin ja David Bourgetin (2014) kyselytutkimuksen mukaan filosofien vaikeaa ongelmaa koskevat intuitiot jakoutuivat hyvin eri tavoin, samoin kuin heidän olettamuksensa muiden filosofien intuitioista (merkittävää siinä mielessä, miten huolellista perustelua tutkija katsoo esittämänsä argumentin vaativan).

Tämän tutkimuksen kannalta olennaista on, että eroavista tietoisuuden teorioista huolimatta monet keskeiset intuitiot ovat silti samanlaiset: jopa Daniel Dennett myöntää, että kokemuksellinen tietoisuus *näyttää* ihan ilmiselvästi joltakin jota ei voi selittää fysikaalisesti (esim. Dennett 1991, 27). Näin ollen selvitettävä kysymys on, ovatko ko. intuitiot evidenssiä jonkin ei-fysikaalisen olemassaolosta vai ”tavallisia”, selitystä kaipaavia ilmiöitä. Jos intuitioiden tyhjentävä luonnontieteellinen selittäminen onnistuu, materialisti voi väittää ratkaisseensa vaikean ongelman.

Viime vuosina Chalmers on itsekin analysoinut perustavien intuitioidensa syntyperää (Chalmers 2014, 2018a). Artikkelissa *The Meta-Problem of Consciousness* hän esitteli kehikon, jota hän kutsui tietoisuuden metaongelmaksi: miten syntyvät ns. ongelma-intuitiot, kuten ”tietoisuus on ei-fysikaalinen” tai ”tietoisuus tuottaa erityistä ensimmäisen persoonan tietoa”.¹ Tämä Chalmersin keskustelunavaus ja Dennettin ja muiden illusionistien kritiikki neuvoo tutkimaan tarkemmin realistien tietoisuuskäsityksen perustavia intuitioita. Tämä juuri on tämän tutkielman aiheena.

¹ Muita tätä intuitio-ongelmaa käsitelleitä ovat ainakin: David Papineau (2002, 2011), Pär Sundström (2008) ja François Kammerer (2018).

Erityisesti tutkielmassa tarkennetaan Dennettin esittämään kritiikkiin ja siitä tehtäviin johtopäätöksiin. Osa intuitioista on osoitettavissa ristiriitaisiksi, osa taas vastustaa argumentointia eikä niitä voi selkeästi *todistaa* virheellisiksi tai perustelluiksi. Jälkimmäisten tapauksessa kuitenkin materialistit löytävät monesti kattavamman tai taloudellisemman *selityksen*, kun taas realisteilla ei useinkaan ole selitystä ollenkaan.

Dennettin kirjoituksia joudutaan lapioidaan laajalti ja lavealti, mutta tämä on perusteltua, sillä hänen kritiikkinsä kohteet ovat kuitenkin selvästi rajattavissa. Koska Dennett systemaattisen Chalmersin argumenttien purkamisen sijaan keskittyy enemmänkin hyökkäämään – verbaalisesti loistavan taitavasti – eri suunnista Chalmersin etukäteisolettamuksia kohtaan, ajatuskokein, empiirisen tieteen tiedonsirpalein ja vertauksin, häntä koskeva osio muodostuu välttämättä vähemmän strukturoiduksi kuin Chalmersin ajattelua esittävä.

Näin rikkaan ja vuosia jatkuneen keskustelun kaikkien argumenttien seuraaminen ei ole tässä mahdollista, mutta pyrin luomaan yleiskatsauksen realistien ja illusionistien ja erityisesti Chalmersin ja Dennettin näkökantoihin, etsimään virheitä Chalmersin metodologiasta ja muodostamaan käsityksen siitä, kumman tietoisuuden selitysmalli vaikuttaa tieteellisen teorianmuodostuksen periaatteet huomioiden hyväksyttävämmältä.

1.2 Tutkielman rakenne

Luvussa 1 kuvailen ensin Chalmersin ongelman ja sen motivaation. Luvussa 2 esittelen materialismin keskeiset käsitteet ja ongelman, jonka fenomenaalinen tietoisuus näyttää muodostavan materialismille. Luku 3 on omistettu argumenteille, jotka voi muotoilla materialismia vastaan jos vaikea ongelma hyväksytään olemassa olevaksi, ja luvussa 4 esittelen Chalmersin tietoisuuden teorian, joka näyttäisi olevan seurauksena mikäli luvun 3 argumentit hyväksytään päteviksi. Luvussa 5 ryhmittelen intuitiot, joihin Chalmersin ja muiden realistien argumentointi tukeutuu.

Luvussa 6 aloitan Chalmersille vastakkaisten näkemysten esittelyn *illusionismin* katto-käsitteen alla. Luvussa 7 esittelen Dennettin heterofenomenologian ja luonnostelen hänen tietoisuusteoriaansa, jonka jälkeen luvussa 8 käsittelen Dennettin realistien intuitioihin kohdistamaa kritiikkiä. Seuraavassa luvussa 9 teen edellisestä yhteenvedon ja arvioin mitä Chalmersin lähestymistavasta jää jäljelle. Päätävässä luvussa 10 etsin yhtymäkohtia Chalmersin ja Dennettin näkemysten välillä ja esitän joitakin arvioita tietoisuuden tutkimuksen tulevista suuntauksista.

2 Taustaa

2.1 Materialismi

Kartesiolaisen dualismin ongelmien takia ja luonnontieteiden menestyksen jälkeen filosofian huomio siirtyi paljolti pois subjektiivisen tietoisuuden tutkimuksesta. Viime vuosikymmeninä keskustelu on taas vilkastunut, kun behaviorismi on väistynyt monialaisen kognitiotieteen tieltä, komputatiiviset teoriat ovat kehittyneet ja tietokoneiden lisääntynyt laskentateho on sallinut entistä monipuolisempien kognitiivisten mallien konstruoimisen. Myös hermoverkko- ja tekoälytutkimus on herättänyt kysymyksen tietoisuuden luonteesta ja mahdollisuudesta erilaisissa realisaatioissa, jonka vastauksella voi olla myös eettisiä seurauksia. Samoin tapaukset, joissa leikkauspotilas on myöhemmin kertonut olleensa leikkauksen ajan tietoinen kannustavat etsimään jonkinlaista ”tietoisuusilmaisinta” (ks. esim. Mashour & LaRock 2008).

Kognitiivisten ja komputatiivisten mallien ontologiana on yleensä ollut monismi: ne selittävät tietoisuuden reduktiiviseksi perustaksi fysikaalisen todellisuuden. Tätä näkemystä kutsutaan *fysikalismiksi* tai *materialismiksi*: kaikki olemassa oleva koostuu fysiikan käsittelemistä olioista. Fysiikka katsotaan yleensä kausaalisesti sulkeutuneeksi, toisin sanoen jokaisella fysikaalisella tapahtumalla on oltava fysikaalinen syy. Tämä on myös metodologinen periaate: fysiikan tutkijoiden ei voi ajatella postuloivan selityksiinsä fysiikan ontologian ulkopuolisia elementtejä – kuten mentaalisia, ei-fysikaalisia – joita ei voisi edes periaatteessa havaita ja joiden kausaalinen rooli ei olisi fysikaalisiin menetelmin tutkittavissa.

Materialismi taas on sitoutunut materian olemassaoloon, ts. kaikki oleva on selitettävissä aineen, aineen liikkeen ja sitä ohjaavien fysikaalisten lakien avulla. Erityisesti mielen filosofian kontekstissa materialismi esittää fysiikasta käsityksen, joka on nykyfysiikan valossa selvästi vanhanaikainen: ”on vain ainehiukkasia ja niiden fysiikan lakien mukaan käyttäytyviä yhdistelmiä” (Kim 2005, 149). Tietoisuuden tutkimuksessa fysikalismin ja materialismin käsitteitä voi kuitenkin käyttää enimmäkseen samassa merkityksessä. Näin teen myös tässä tutkielmassa jatkossa.

Fysikalistinen tietoisuuden selitys on yleensä *reduktiivinen*: tietoisuuden ominaisuudet ovat täydellisesti, jäännöksettä selitettävissä fysiikan postuloimien entiteettien ja lakien avulla. (Chalmers 2002a, 248) Tällöin löytyy jokaista tietoisuuden ominaisuutta vastaava fysikaalinen ominaisuus, ja näiden välille voi asettaa identiteetin. Jos ilmiö voi realisoitua

erilaisissa fysikaalisissa järjestelmissä, joudutaan tyytymään *reduktiiviseen selitykseen* – joka on tavallisesti *funktionaalinen* (Chalmers 1996, 43). *Ei-reduktiivinen* fysikalismi perustuu yleensä supervenienssiin², jonka mukaan mentaaliset ominaisuudet määräytyvät fysikaalisten ominaisuuksien perusteella, mutta ominaisuuksien välillä ei ole identiteettiä. Funktionalistiset mielen teoriat ovat tavallisesti supervenienssiteorioita tässä mielessä.

Funktionalistisessa mielen teoriassa mentaaliset tilat määritellään niiden *kausaalisten roolien* avulla. Jokaista mentaalista ominaisuutta vastaa jokin fysikaalinen ominaisuus joka toteuttaa eli *realisoi* mentaalisen ominaisuuden, ja näitä realisoijia voi olla erilaisia. Mentaalisilla tiloilla ajatellaan olevan todellisia sisäisiä tiloja omine kausaalisine rooleineen, eikä vain instrumentalistisesti ajateltuja kuten behaviorismissa. *Reduktiivinen funktionalismi* sopii yhteen materialistisen käsityksen kanssa: riittää selittää, miten psykologisten tilojen kausaaliset roolit on realisoitu, esimerkiksi esittämällä niiden perustana oleva neurofysiologia tai rakentamalla sopiva kognitiivinen malli.

Chalmers (1996, 2002a) erottaa toisistaan A-tyypin ja B-tyypin³ materialistit. A-tyypin materialistit kieltävät episteemisen kuilun olemassaolon, eli heille mitään selitettävää vaikeaa ongelmaa ei ole. Huomattavimpia ja jyrkimpiä A-tyypin materialisteja on Dennett, joka on suoraan sanonut, että hänen introspektionensa ei havaitse mitään fenomenaalisia tiloja funktioiden, struktuurien ja fenomenaalisten arvostelmien lisäksi, ”pelkääntään pimeyttä” (Dennett 1979, 95). Samoin Dennett on kritisoinut käsitystä kvalioista intrinsisinä, rakenteettomina, episteemisesti ensisijaisina ja yksityisinä (esim. Dennett 1988).

B-tyypin materialistit hyväksyvät episteemisen kuilun olemassaolon, mutta kieltävät ontologisen kuilun olemassaolon. He hyväksyvät vaikean ongelman olemassaolon, mutta uskovat, että tarkempi analyysi pystyy selittämään sen pois: johtuu vain tietoisuutta ja vastaavasti fysikaalista koskevien käsitteidemme suhteesta, ettemme näe yhteyttä kokemuksellisen ominaisuuden ja fysikaalisten tosiasioiden välillä.⁴

² Supervenienssin tarkempi määritelmä löytyy esimerkiksi Jaegwon Kimiltä (2005, 14).

³ Tilarajoitusten vuoksi B-tyypin materialismin käsittely on tässä tutkielmassa jätettävä muutamaan haja-huomioon.

⁴ Samoin sivuutetaan Chalmersin mainitsema C-tyypin materialismi, jonka mukaan episteeminen kuilu on olemassa, mutta se on periaatteessa suljettavissa. Näkemyksessä useimmiten oletetaan, että kuilun ylittämisen vaikeus johtuu kognitiivisista rajoitteistamme. Chalmersin mielestä C-materialistinen kanta ei kuitenkaan kestä vaan romahtaa A- tai B-tyypin materialismiin, interaktionismiin tai russelilaiseen monismiin. (Chalmers, 2002a, 258–260.)

2.2 Tietoisuuden muodostama ongelma

Chalmersin mukaan, vaikka tietäisimme kaiken tiedettävissä olevan fysikaalisesta maailmasta, mikään tässä tiedossa ei saisi meitä postuloimaan tietoisuuden kokemuksen olemassaoloa – vain se, että meillä on omaa kokemusta siitä. Ensimmäisen persoonan kokemus tietoisuudesta pakottaa ongelman esiin. (Chalmers 1996, 101–2.) Kaikista fysiikan matalan tason tosiasioista ja laeista pystymme johtamaan kemialliset ja biologiset ilmiöt, kognitiiviset prosessit ja aivokemian, mutta tietoisuus on *yllättävä* ilmiö (Chalmers 1996, 5), jotakin, jonka olemassaoloon emme osaisi uskoa ilman omaa kokemustamme siitä.

Tietoisuuden erityisluonteeseen ovat ennen Chalmersia kiinnittäneet huomiota esimerkiksi Thomas Nagel, Joseph Levine ja Frank Jackson. Nagel (1974) kysyi, miten meillä voisi olla tietoa toisen olennon kokemuksesta, siitä *millaista on olla* tuo olento? Levine argumentoi, että fysikaalisten prosessien ja tietoisuuden välillä on *selityksellinen kuilu*, jota ei voi ylittää millään fysikaalisella mallilla (Levine 1983, 1994; ks. myös Ned Block 1995, 231). Jackson (1982) esitti ajatuskokeen mustavalkoisessa huoneessa asuvasta väritutkija Marysta, joka oletuksen mukaan tietää kaiken fysikaalisen mitä väreistä voi tietää mutta (intuition mukaan) vaikuttaa kuitenkin oppivan jotakin uutta päästessään ulos ja nähdessään värejä ensi kertaa. Chalmersin jaottelu fenomenaaliseen ja psykologiseen tietoisuuteen vastaa myös Blockin (1995) erottelua *fenomenaaliseen tietoisuuteen* (phenomenal consciousness) ja *pääsy tietoisuuteen* (access consciousness)⁵. Chalmersin vaikea ongelma ei ole Descartesin mieli-ruumis-ongelma, jota Chalmers pitää ratkaistuna – psykologiset tilat syntyvät aivotiloista – vaan yhteys psykologisen tietoisuuden ja fenomenaalisen tietoisuuden välillä (Chalmers 1996, 25).

Fenomenaalisen tietoisuuden keskeisiä piirteitä ovat sen laadulliset ominaisuudet. Esimerkiksi jonkin seinässä olevan punaisen läiskän heijastamasta valosta voidaan mitata sen spektri, analysoida välillä olevan ilman vaikutus, valon saapuminen silmään, sauvojen ja tappien toiminta, ihmisen värierotteluvaruuden puna-viher- ja sini-kelta-erottelufunktioiden vaikutus, ja viimein erottelukyvyn rajoissa seurata havainnon vaikutuksia näkökeskuksessa. Tämä kaikki voidaan kuvata funktionaalisesti. Lisäksi voidaan kuulla subjektin suullinen raportti, jonka sisältö on vaikkapa ”Näen punaista” tai ”Tämä on kirkkaampi punainen kuin tuon kirjan kannessa”. Mutta näin saadaan vain funktionaalinen selitys punaisen värin näkemiselle, mikään tässä ei selitä, miksi siihen

⁵ Tarkemmin: Blockin pääsy tietoisuus vastaa Chalmersin *a-tietoisuutta*, ks. luku 4.2.

liittyy lisäksi *kokemus* punaisen aistimisesta. Voisihan olla, että kaikki tämä tapahtuisi ilman minkäänlaista kokemusta. Tai onko varmuutta, että väriaistimukset ovat samanlaiset eri ihmisillä: ehkä joku voi kokea sinisenä värin jota minä sanoisin punaiseksi?

Chalmersin mielestä kolmannen persoonan havaintoaineiston lisäksi myös ensimmäisen persoonan aineisto vaatii selityksen. Esimerkkinä musiikki: on varsin ilmeistä, että kolmannen persoonan havainnot, kuten huomiot henkilön käyttäytymisestä, kuuloluiden toiminnasta, fMRI-skannaukset kuulokeskuksen prosesseista, eivät riitä selittämään musiikkia. Pelkästään kolmannen persoonan aineistoon perustuva musiikin tiede olisi hyvin puutteellista. (Chalmers 2004a, 1111–2.)

Näin Chalmers argumentoi *analyysin puuttumisen* todisteeksi reduktiivisen fysikalismin epäonnistumiselle: fysikalisteilla on tarjolla vain funktionaalinen analyysi kokemuksellisen tietoisuuden synnystä – mikä ei riitä, sillä vaikka tietoiset tilat esiintyvät monissa kausaalisissa rooleissa, nämä roolit eivät määrittele tietoisia tiloja (Chalmers 1996, 104). Kuitenkin, kuten jäljempänä nähdään, fysikalisteilla näyttää sittenkin olevan uskottavia analyyseja vaikean ongelman syntyyn.

3 Argumentteja materialismia vastaan

Chalmersin mukaan eliminativismi tietoisuuden suhteen ei siis vaikuta järkevältä koska meillä on omakohtainen kokemus tietoisuudesta. Chalmersin myöntää, että kaikki tässä luvussa esitettävät argumentit perustuvat tähän pääintuition, jonka mukaan ensimmäisen persoonan kokemuksessa *on jotakin joka vaatii selitystä* (Chalmers 1996, 110). Tiedossamme tietoisuudesta on *episteeminen epäsymmetria*: mikään kausaalisessa tarinassa fysikaalisista faktoista biologiaan ja psykologiaan ei saisi uskomaan tietoisuuden olemassaoloon – paitsi henkilön, joka on kokenut sen itse. Tällaista epäsymmetriaa ei ole tiedossamme muista ilmiöistä: ei ole minkäänlaista ”toisten elämien ongelmaa” tai ”toisten talousjärjestelmien ongelmaa”, sillä nämä supervenioivat loogisesti fysikaalisen yllä. (Chalmers 1996, 102.)

Chalmers määrittelee *loogisen supervenienssin* näin: ylempi taso supervenioi loogisesti alemman tason yllä jos ei ole mahdollista, että kaksi loogisesti mahdollista tilannetta ovat alemman tason ominaisuuksiltaan identtiset mutta eroavat ylemmän tason ominaisuuksiltaan (Chalmers 1996, 34). *Luonnollisessa supervenienssissa* vastaavasti ei ole mahdollista, että kaksi *luonnollisesti* (empiirisesti) *mahdollista* tilannetta ovat alemman

tason ominaisuuksiltaan identtiset mutta eroavat ylemmän tason ominaisuuksiltaan.⁶

Chalmersin mukaan jo se, että filosofiassa on olemassa ”toisten tietoisuuksien ongelma” todistaa, että tietoisuus ei voi supervenioida loogisesti fysikaalisen yllä (Chalmers 1996, 74). Ja aivan keskeistä Chalmersin argumentoinnille on, että jos tietoisuus ei supervenoi loogisesti fysikaalista, tietoisuuden reduktiivinen selitys ole mahdollinen.

Luonnostelen aluksi Chalmersin kaksiulotteisen semantiikan, jota Chalmers käyttää usein tukemaan modaalisia argumenttejaan (kuten luvussa 3.4 esitettävää tietoargumenttia). Sen jälkeen esittelen fysikalismia vastaan suunnatun *selitysargumentin*, jonka Chalmers muotoilee Nageliin ja Levineen tukeutuen, Chalmersille keskeisen *kuviteltavuusargumentin*, sekä Jacksonin *tietoargumentin*, jota Chalmers niin ikään käyttää fysikalismien rajoitusten osoittamiseen. Chalmers kutsuu näitä kaikkia *episteemiseksi argumenteiksi* materialismia vastaan, sillä ne kaikki rakentavat episteemisen kuilun fysikaalisen ja fenomenaalisen välillä (Chalmers 2002a, 250). Argumentit kieltävät, että fysikaalisista totuuksista P voi dedusoida fenomenaalisia totuuksia Q, että P:n avulla voi selittää Q:n, tai että P:n kuviteltavuudesta seuraisi Q:n kuviteltavuus. Tämän jälkeen argumenteissa johdetaan päätelmä ontologisesta kuilusta: selitysargumentissa päätellään fysikaalisen selityksen pettämisestä ei-fysikaalisten ominaisuuksien olemassaolo, kuviteltavuusargumentissa päätellään kuviteltavuudesta metafyyssinen mahdollisuus, ja tietoargumentissa päätellään dedusoinnin epäonnistumisesta että tietoisuuden ominaisuuksia ei voi johtaa fysikaalisista ominaisuuksista.

3.1 Kaksiulotteinen semantiikka

Chalmers käytti kaksiulotteista semantiikkaa ensimmäisen kerran, kun hän käsitteli Saul Kripken (1972) väitettä aposterioristen välttämättömyyksien (kuten ”vesi = H₂O”) olemassaolosta (Chalmers 1996, 56). Chalmersin määrittelyssä *käsite* määrittelee funktion $f : W \rightarrow R$ mahdollisilta maailmoilta referenteille; tämä funktio on käsitteen *intensio*, joka yhdessä maailman $w \in W$ kanssa määrittelee käsitteen *ekstension* $f(w)$. Chalmersin mukaan intensioita tarvitaan kuitenkin kaksi.

Primaari (eli *episteeminen*) *intensio* on funktio, joka esittää mikä käsitteen referentti maailmassa w olisi, jos w olisi *aktuaalinen*, ts. jos w sattuisi olemaan *oma* maailmamme

⁶ Esimerkiksi kaasun paine supervenoi luonnollisesti kaasun lämpötilan ja tilavuuden yllä lain $pV = KT$ mukaan, vaikka onkin kuviteltavissa loogisesti mahdollinen maailma, jossa vaikkapa vakiolla K olisi eri arvo.

(esim. ”veden” primaarin intension mukainen referentti on ”hallitseva kirkas, juotava neste ympäristössämme”, on veden molekyyli rakenne sitten H₂O tai vaikkapa XYZ, jos maailmamme sattuisi olemaan Putnamin (1975) Kaksois-Maa). Väite S on *episteemisesti mahdollinen*, jos se ei ole a priori mahdoton (on esimerkiksi episteemisesti mahdollista, että ”vesi ei ole H₂O:ta”, koska on episteemisesti mahdollista että oma maailmamme on Kaksois-Maa). Jotta episteeminen mahdollisuus olisi määritelty, maailman pitää olla aktuaalinen ja *keskitetty*: kiinnitetty tiettyyn yksilöön ja ajankohtaan (Chalmers 1996, 60; 2004b, 295).

Sekundaarinen (eli *subjunktiivinen*) *intensio* on funktio, joka poimii referentin kaikissa *kontrafaktuaaleissa* maailmoissa, kun primaarin intension kiinnittämästä aineesta on empiirisesti selvitetty kaikki tiedettävissä oleva (jos ”vesi” sattuu olemaan H₂O:ta aktuaalisessa maailmassa, sekundaarinen intensio poimii ”veden” referentiksi H₂O:n kaikissa kontrafaktuaaleissa maailmoissa, eikä ole enää kuviteltavissa, että ”vesi ei ole H₂O:ta”). Chalmersin määritelmien mukaan väite S on *1-mahdollinen* (*primaaristi mahdollinen*) jos sen episteeminen intensio on tosi jossakin maailmassa (käsitettynä aktuaalisena). Väite S on *1-kontingenti*, jos sen episteeminen intensio on epätosi jossakin maailmassa (käsitettynä aktuaalisena). Väite S on *2-mahdollinen* (*sekundaarisesti mahdollinen*) jos sen subjunktiivinen intensio on tosi jossain maailmassa (käsitettynä kontrafaktuaalisena). Väite S on *2-kontingenti*, jos sen sekundaarinen intensio on epätosi jossakin maailmassa. (Chalmers 2004b, 281 ja 295.) Primaarinen ja sekundaarinen *välttämättömyys* määritellään vastaavasti.

Tavallisemmalla käsitteistöllä, väite S on *loogisesti mahdollinen*, jos se on tosi jossain maailmassa primaarin intension mukaan, ja väite S on *metafyysisesti mahdollinen*, jos se on tosi jossain maailmassa sekundaarin intension mukaan.

3.2 Selitysargumentti

Nagelin uupumukseen asti siteeratussa artikkelissa (Nagel 1974) hän kysyy, miten voisimme käsittää lepakon kokemuksellista tietoisuutta, sitä millaista on olla lepakko⁷. Nagel toteaa mahdottomaksi käsittää lepakon kokemuksellista tietoisuutta, miten paljon

⁷ Nagel valitsi tarkoituksella esimerkikseen olennon, joka on riittävän erilainen meihin verrattuna (sillä vaikkapa kissan tai koiran kokemusta voisimme jotenkin kuvitella ymmärtävämmeksi) mutta ei liian erilainen (kärpäsän tapauksessa alkaisimme jo epäillä, voiko sillä olla ollenkaan kokemuksellista tietoisuutta). Myös Searlilla on samantapainen argumentti, jossa hän kehottaa yrittämään kuvitella, millaista olisi olla lintu ja tuntea, miten Maan magneettikenttä ohjaa lentämistä (Searle 1992, 72).

tahansa tietoa meillä onkin lepakon fysiologiasta, aistijärjestelmistä tai kognitiivisista prosesseista. Kaikki tieto mitä voimme lepakosta saada, on ilmaistava kolmannen persoonan neutraalilla ja fysikalistiselle tieteelle sopivalla kielellä, jonka on periaatteessakin mahdotonta kuvata lepakon subjektiivista kokemusta. Paitsi ettemme pääse objektiivisesti käsiksi lepakon subjektiiviseen kokemukseen, emme pääse siihen käsiksi myöskään subjektiivisesti. Voisin kuvitella itselleni kaikuluotaimen, että roikkuisin päivisin jonkin luolan katossa ja lentäisin pimeällä pyydystämässä hyönteisiä, mutta näin sain tietää vain millaista olisi olla *minä*, joka käyttäydyn kuten lepakko. Toisin sanoen, fysikaalisin käsittein ei voi kuvata tietoisuuden jonkinlaisuutta, mikä kuitenkin on juuri olennaisin selitettävä ilmiö. Joten fysikalismi ei siis objektiivisella kielellään periaatteesakaan voi kuvata kokemuksellisuutta (Nagel 1974, 445). Nagelin argumentti on episteeminen: hän haluaa tuoda esiin, miten meillä ei ole alustavaakaan käsitystä siitä, miten fysikalismi voisi selittää subjektiivisen tietoisuuden, mutta hän ei käytä argumenttia väittääkseen fysikalismin olevan väärässä.

Levine puolestaan koettaa osoittaa, että tietyt väitteet identiteeteistä mentaalisten ja fysikaalisten ominaisuuksien välillä jättävät selityksellisen kuilun joka tuottaa vaikeuksia fysikalismille. Hän tekee tämän rakentamalla epistemologisen väitteen Kripken psykofyysistä identiteettiteoriaa kohtaan suuntaamasta kritiikistä. Kripken analyysin mukaan, jos molemmat puolet identiteettiä ovat jäykkiä nimittäjiä, jos identiteetti on tosi se on tosi välttämättä (Kripke 1972, 108). Siis jos psykofyysinen identiteetti kuten ”kipu = C-hermosäikeiden laukeaminen”⁸ on tosi se on välttämättä tosi. Mutta koska tämä identiteetti on helposti kuviteltavissa epätodeksi, psykofyysiset identiteetit eivät näytä olevan välttämättä tosia, ja siis tosia ollenkaan. Vastaväite, että tämä identiteetin kontingenssi on vain näennäinen ja edellisessä esimerkissä kivun tuntemus johtuu muusta kuin kivusta ei toimi, koska – Kripken intuition mukaan – tällaisten kvalitatiivisten kokemusten tapauksessa näennäinen on sama asia kuin itse ilmiö (Kripke, 1972, 152–4). Joten psykofyysinen identiteettiteoria ei voi olla tosi.

Levine käyttää esimerkkeinä Kripken esittämiä tyyppi-identiteettiteorian mukaisia väitteitä:

- (1) Kipu = C-hermosäikeiden laukeaminen (Cfs).
- (2) Lämpö = molekyylien liikettä.

⁸ ”Filosofien fiktiivisessä neurotieteessä”, kuten Kim sanoo (Kim 2005, 13).

Kummassakin identiteetissä on ”tuntuva kontingenssi” (Levine 1983, 355), ts. on ajateltavissa, että ne ovat epätosia. Identiteetin (2) kontingenssi on kuitenkin selitettävissä siten, että voimme kuvitella jonkin toisen ilmiön joka vaikuttaa aisteihimme kuten lämpö, mutta ei ole molekyylien liikettä. Identiteetin (1) tapauksessa tämä selitys ei toimi, sillä jos kuvittelemme maailman jossa jokin ilmiö koetaan kipuna mutta siellä ei ole Cfs:ää, kuvittelemme maailman jossa on kipua mutta ei Cfs:ää (Levine kuten realistik yleensäkin nojaavat tässä samaan intuitioon kuin Kripke). Nyt materialisti voi koettaa ottaa funktionalistisen lähestymistavan, jolloin väitteen (1) kontingenssi näyttää monitoteutuvuuden vuoksi ongelmattomalta. Tähän Levine väittää, että Kripken argumentti toimii samalla tavalla funktionalismia vastaan. Funktionalistisesti

(3) Kokea kipua = olla tilassa F.

On taaskin hyvin kuviteltavissa, että (3) ei ole tosi, esimerkkinä Blockin (1978) ajatuskoe, jossa F on realisoitu Kiinan kansan funktionaalisena organisaationa, jolloin on luontevaa kuvitella että tämä realisaatio ei koe kipua. Levinen mukaan materialistin täytyy nyt intuitionsa vastaisesti kieltää, että (1) tai (3) ovat kontingenteja, tai luopua materialismistaan.

Identiteetti (2) on *täydellisesti selitettävissä*, kun taas väittämissä (1) ja (3) näyttää jäävän jotain selittämättä. Tätä Levine kutsuu *selitykselliseksi kuiluksi*. Jos emme löydä C-hermosäikeiden laukeamisesta mitään, mikä selittäisi miksi Cfs:llä on se laadullinen luonne mikä sillä on, on heti kuviteltavissa, että Cfs voisi esiintyä ilman kivun tuntemusta tai päinvastoin. Vastaavaa intuitiota ei synny tapauksessa (2). (Levine 1983, 359.)

Lisäksi, vaikka jotkin psykofysikaaliset identiteetit olisivatkin tosia, emme pystyisi määrittelemään mitkä niistä ovat tosia (Levine 1983, 360)! Vaikka kivun tuntemus olisikin fysikaalinen tila, meillä ei olisi mitään keinoa määritellä, minkä fysikaalisten ominaisuuksien kanssa tarkalleen se olisi identtinen. Ja samaa argumenttia voi soveltaa funktionalistisiin identiteettiväitteisiin: vaikka kivun tuntemus olisikin identtinen jonkin funktionaalisen tilan kanssa, meillä ei olisi keinoa arvioida, mitkä funktionaaliset samantyyppiset tai eroavaisuudet olisivat olennaisia arvioidessamme esimerkiksi ulkoavaruuden olennon kykyä tuntea laadullisesti samanlaista kipua.

Toisin sanoen, väitteen (1) totuus on *epistemologisesti saavuttamattomissa* ihan riippumatta siitä onko se metafyyminen totuus (Levine 1983, 360).

Levine tunnustautuu ”kvaliofiiliksi” (erotuksena ”kvaliofoobikosta”), eli hän uskoo, että kokemuksellinen tietoisuus muodostaa haasteen materialismille. Hän ei kuitenkaan ole

valmis väittämään, että oman tietoisien kokemuksen havaitsemisesta voi päätellä sen materialistisen selityksen mahdottomuuden – vain että kokemuksellinen tietoisuus muodostaa erikoisen haasteen materialismille. (Levine 1994, 107.) Sen sijaan Chalmers on valmis muodostamaan selityksellisen kuilun perusteella ontologisen väitteen.

Chalmersin luokittelussa helpot ongelmat ovat niitä jotka voi selittää funktionaalisesti esittämällä mekanismit jotka suorittavat funktion (Chalmers 1995, 202). Siis esittämällä esimerkiksi hermostolliset mekanismit, jotka ovat vastuussa tarkastelemastamme ilmiöstä, tai abstraktimman kuvauksen vaikkapa komputatiivisin käsittein. Esimerkiksi jotta järjestelmä kykenee kuvailemaan mielentilojaan, riittää että järjestelmällä on kyky verbaalisiin raportteihin ja pääsy sisäisiä tiloja koskevaan informaatioon; samalla tavalla reduktiivisesti voidaan selittää geeni, elämä, oppiminen tai muisti (Chalmers 1995, 203). Kun tällaisen ilmiön koko kausaalinen organisaatio on selitetty, muuta selitettävää ei enää jää. Kun vaikkapa oppimisen funktionaalinen kuvaus on annettu, on *loogisesti mahdotonta*, että nuo funktiot instantioituisivat mutta oppimista ei tapahtuisi.

Mutta fenomenaalisia tiloja eivät määrittele niiden kausaaliset roolit. Vaikka selittäisimme kaiken minkä voimme funktionaalisesti selittää, jää aina silti *lisäkysymys*: miksi näiden funktioiden toimintaan liittyy myös tietoinen kokemus (Chalmers 1995, 203; 1996, 47)? Ja näyttää loogisesti mahdolliselta, että minkä tahansa funktionaalisen kuvauksen kognitiosta esittäisimme, kyseiset funktiot voisivat instantioitua ilman niihin liittyvää tietoisuutta (Chalmers 1996, 47). Chalmers lainaa Levinen käsitettä ja sanoo, että funktionaalisten ja fysikalististen selitysten ja tietoisuuden välillä on selityksellinen kuilu. Chalmersin etsimien psykofyysisten lakien tehtävänä olisi silloittaa tämä kuilu. Hänen radikaalein ja kiistanalaisin lisäys on kuitenkin ehdottaa edeltävän perusteella myös, että kokemuksellisuus tulisi ottaa perusilmiönä, samaan tapaan kuin massa tai sähkövaraus (Chalmers, 1995, 210)⁹.

3.3 Kuviteltavuusargumentti

Chalmersin kuviteltavuusargumentilla¹⁰, erityisesti zombiargumentilla on edeltäjiä: Keith Campbell (1970) esitti ajatuksen ”imitaatiomiehestä” todistellakseen psykofyysistä

⁹ Kokemuksellisuus samoin kuin psykofyysiset lait olisivat siis vain raakoja, kontingenteja tosiasioita. Mark Pricen (1996) mielenkiintoisen ajatuksen mukaan tällaiset raat kontingenssit ovat itse asiassa tavallisia. Samanlaisia selityksellisiä kuiluja on luonnontieteissä joka puolella, esimerkiksi syy-seuraussuhteissa – me vain olemme totuneet niihin.

¹⁰ Tästä käytetään myös nimitystä modaaliargumentti. Zombiargumentti taas on kuviteltavuusargumentin tietty muoto, ns. puuttuvan kvalian argumentti.

identiteettiteoriaa vastaan, ja Nagel (1970, 1974) ja Robert Kirk (1974) rakensivat zombi-idean avulla fysikalismiin vastaisia argumentteja. Myös Jackson (1982) esitteli modaaliargumentin lyhyesti, John Searle argumentoi että kaksi järjestelmää joista toinen on tietoinen ja toinen ei-tietoinen voivat käyttäytyä identtisesti (Searle 1992, 70), ja samoin Block piti käsitteellisesti mahdollisena, että pääsytietoisuus voi esiintyä ilman fenomenaalista tietoisuutta ja päinvastoin (Block 1995, 233).

Zombit tässä ovat siis nk. filosofisia zombeja, ulkoisesti täysin kaltaisiamme eivätkä Hollywood-elokuvien katalasti deformatuneita ja selvistä funktionaalisista vioista kärsiviä olentoja. Argumentti ei ole ilahduttanut niitä joiden mielestä zombin käsitteen olisi syytä saattaa koko ammattikunta kiusaantuneeksi (Dennett 1995). Ajatuskoe on joka tapauksessa ilmeikäs, ja esitän sen seuraavassa jälleen kerran.

Kuvitellaan meidän maailmamme kanssa fysikaalisesti identtinen maailma, jossa asuu minun fysikaalisesti identtinen kaksoisolentoni. Zombiekaksoseni käyttäytyy täsmälleen kuten minä, syö, opiskelee ja nukkuu ja päivittelee, miten hankala tietoisuuden muodostama filosofinen ongelma on. Onnettomasti vain tällä kaksoisolennolla ei ole lainkaan fenomenaalista tietoisuutta, ts. ei ole mitään, millaista olisi olla zombi. Kun zombiekaksoseni lukee artikkelia paperilta, se osaa erottaa funktionaalisesti painomusteen valkoisesta paperista, jotta sen kognitiivinen järjestelmä kykenisi oppimaan, mutta sillä ei ole mitään kokemuksellista käsitystä paperin valkoisuudesta. Jos se potkaisisi varpaansa kiveen, se käyttäytyisi kuten minä, ääntelisi samaan tapaan, mutta sillä ei olisi kokemusta kivusta. Kaikki olisi pimeää sisällä.

Tässä kuvitelmassa ei näytä olevan mitään epäkoherenttia. Me voimme ainakin prima facie kuvitella zombien ristiriidattomasti, Chalmers sanoo. Levine sanoin tämä kuviteltavuus käy ilmeiseksi siitä, että aikoinaan tarvittiin empiirinen havainto sen toteamiseksi, että ajattelemme ja koemme aivoillamme emmekä sydämillämme (Levine 2009, 283). Vaikkakaan filosofiset zombiet tuskin ovat empiirisesti mahdollisia, vaikuttaa siltä, että ne ovat loogisesti mahdollisia (Chalmers 1996, 96). Ja tästä seuraa, että tietoisuus ei supervenoi loogisesti fysikaalisen yllä.

Argumentin kannalta ei ole olennaista, kuvitellaanko kokonainen zombi maailma vai vain yksi yksilö, tai kuvittelemmeko tähän maailmaan kanssani funktionaalisesti identtisen puolijohteista tai radiolähettimillä varustetusta Kiinan kansasta rakennetun isomorfin. Ei ole olennaista, onko luultavaa että ko. isomorfilla on *todella* kokemuksellista

tietoisuutta – riittää, että sen puuttuminen on loogisen koherentisti kuviteltavissa. Samoin argumentille riittää, jos on kuviteltavissa, että osa kokemuksista on vaihtunut toisiin, tai että vain osa niistä puuttuu (Chalmers 2002a, 249).

Jo Jackson (1982) havaitsi erään kuviteltavuusargumentin heikkouden. Kuten hän sanoo, kvaliofiileillä on seuraava argumentti: ”Mikään fysikaalinen ei onnistu kuvaamaan ruusun tuoksua, joten fysikalismi on väärässä”. Tämä ei kuitenkaan ole riittävän vakuuttavaa. Modaaliargumentti tukeutuu kiistanalaiseen intuitioon; ”kannattajien laskeminen on keho tapa lähestyä modaaliargumenttia, mutta useinkaan emme pysty parempaan kun on kyse modaali-intuitioista” (Jackson 1982, 131). Artikkelissa Jackson etsii vakuuttavampaa argumenttia, jonka hän sitten löytääkin tietoargumentista.

Kvalioiden inversio

Chalmersin mukaan vaikean ongelman osoittamiseksi ei tarvitse kuvitella zombeja, riittää osoittaa, että *jokin* tietoisuuden kvalitatiivinen tosiasia voi olla eri tavoin kahdella funktionaalisesti identtisellä olennolla, samalla kun näitä koskevat fysikaaliset tosiasiat ovat samat (Chalmers 1996, 99–100).

Chalmersin esittämä invertoitujen kvalioiden argumentti on seuraava. Kuvitellaan, että on olemassa minun kanssani funktionaalisesti ja fysikaalisesti identtinen olento, vaikkapa jossakin maailmamme kanssa identtisessä kaksoismaailmassa. On hyvin kuviteltavissa, että kun minä näen verta, minulla on kokemus punaisen aistimuksesta, mutta kaksois-olennollani on kokemus joka vastaa minun kokemustani sinisen aistimuksesta. Mikään fysikaalisessa tarinassa ei kiellä tämän mahdollisuutta. Tämä on periaatteessa jopa järjestettävissä keinotekoisesti (Broackes 2007, 162). Joten on olemassa jotakin tietoisuuteen liittyvää, jota fysikaalinen ei selitä.

Usein tämä kiistetään esittämällä, että ihmisen väriavaruus kuten myös aistimukset värien lämpöisyydestä ja viileydestä (joilla on funktionaalisia rooleja) ovat sikäli epäsymmetriset, että tällainen inversio johtaisi outoihin tilanteisiin, joissa invertoitu väri ei vastaa funktionaalisesti koettua, fenomenaalista väriä (Broackes 2007, 171). Chalmersin mukaan kuitenkin riittää kuvitella jokin luonnollisesti mahdollinen laji, joka kykenee erottamaan vain kaksi väriä A ja B. Nyt voimme hyvin kuvitella tällaisen lajin edustajan ja sen kaksoisolennon, joilla kokemukset väreistä A ja B ovat vaihtuneet, ja tämä riittää argumentille. Koska kvalioiden inversio on kuviteltavissa näille kahdelle luonnollisesti mahdolliselle ja fysikaalisesti identtiselle yksilölle, fysikalismi ei voi selittää kaikkea

kokemuksellisesta tietoisuudesta. (Chalmers 1996, 101.)

Tämän argumentin pitävyys riippuu kuitenkin kvalioiden luonteesta: onko intuitiivinen käsityksemme niistä oikea?

3.4 Tietoargumentti

Jackson (1982) argumentoi, ettei kvalioita voi redusoida fysikaaliseen tietoon. Ajatusko-
keessa hän esittelee Fredin, jonka värinäkö on selvästi parempi kuin meidän: Fred
kykenee erottamaan kaksi erilaista punaisen sävyä, jotka meistä näyttävät täysin saman-
laisilta. Me muut haluaisimme tietää, miten Fred kokee nämä värit, mutta mikään fysi-
kaalinen tieto Fredin näköjärjestelmästä, hermostosta tai käyttäytymisestä ei auta meitä
tietämään sitä. Joten fysikalismi ei voi selittää tietoisuutta (Jackson 1982, 128–30.)

Jacksonin tunnetumpi ajatuskoe esittelee värien tutkija Maryn (Jackson 1982, 1986), joka
tietää kaiken fysikaalisen mitä väreistä on tiedettävissä: värien aallonpituudet, taittumisen
ja heijastumisen, silmän sauvojen ja tappien ja näkökeskuksen toiminnan, näkökyvyn
oppimisen ja niin edelleen. Mutta Marylla on ikävä rajoitus: hän joutuu tutkimaan
maailmaa mustavalkoisesta huoneesta käsin, ainoana ikkunana ulkomaailmaan mustaval-
koinen televisio. Hänellä ei siis ole lainkaan kokemusta värien aistimisesta. Eräänä
päivänä Mary pääsee ulos huoneestaan ja näkee ensimmäistä kertaa värejä (vaihtoehtoi-
sessa versiossa Mary on värisokea ja saa sittemmin leikkauksen ansiosta normaalin
värinäön). Jacksonin kysymys on, oppiiko Mary nyt jotakin uutta?

Vaikuttaa siltä, että Mary oppii uuden tosiasian (”ahaa, punaisen kokemus on siis
tällainen”), joka ei kuulunut hänen täydelliseen fysikaaliseen tietoonsa. Tästä Jackson
päättelee, että fysikaalinen tieto ei sisällä kokemuksellista tietoa väreistä ja fysikalismi ei
voi reduktiivisesti selittää tietoisuutta (Jackson 1982, 130).

Edellä esitetty on yleinen tapa muotoilla tietoargumentti, mutta Chalmersin mielestä liian
suoraviivainen eikä aivan pitävä (Chalmers 2004b, 287; 2009, 192).¹¹ Jacksonin
argumentissa ”fysikaaliset tosiasiat” ovat välillä *kapeasti määriteltyjä* fysikaalisia
tosiasioita jostakin rajoitetusta alueesta kuten mikrofysiikasta (f-tosiasiat) ja välillä
laveasti määriteltyjä fysikaalisia tosiasioita joihin sisällytetään korkeamman tason
fysikaalisia tosiasioita, jotka supervenioivat kapeasti määriteltyjen fysikaalisten faktojen
yllä (F-tosiasiat). Jacksonin argumentissa esitetään, että Mary tietää kaikki f-tosiasiat,

¹¹ Jackson itse kiisti myöhemmin tietoargumentin pitävyyden (Jackson 2003).

mutta materialismille riittää, että fenomenaaliset tosiasiat ovat F-tosiasioita. Joten jos tietoargumentissa ”fysikaaliset tosiasiat” viittaavat f-tosiasioihin, johtopäätös ei kumoa materialismia. Ja jos ”fysikaaliset tosiasiat” viittaavat F-tosiasioihin, premissi jonka mukaan Mary tietää kaikki fysikaaliset faktat ei voi olla tosi. (Chalmers 2009, 192–3.)

Chalmers muotoilee argumentin kaksiulotteisen semantiikan avulla seuraavasti. Olkoon P täydellinen mikrofysikaalinen tieto maailmasta, ja Q puhdasta fenomenaalista käsitettä käyttävä väite¹², jonka mukaan Marylla on fenomenaalisen punaisen kokemus. Lähtökohdainen intuitio on siis, että ei ole a priori tiedettävissä, että $P \rightarrow Q$ (zombiargumentin ja invertoidun spektrin argumentin perusteella väite $P \rightarrow \sim Q$ ei johda ristiriitaan¹³). Tämä on episteeminen väite, josta on vielä päästävä metafyyssiseen väitteeseen. (Chalmers 2004b, 279–80.)

Chalmers esittää, että jos $P \rightarrow Q$ on a posteriori totuus, niin $P \rightarrow Q$ on 1-kontingenti (eli on olemassa jokin maailma, jossa väitteen $P \rightarrow Q$ episteeminen intensio on epätosi; esim. väitteen ”vesi = H₂O” episteeminen intensio on epätosi Kaksois-Maa-keskisessä maailmassa) (Chalmers 2004b, 281). Tämä väite mahdollistaa Chalmersin mukaan päättelyn episteemisistä väitteistä väitteisiin metafyyssisestä mahdollisuudesta. Tietoargumentti on nyt (olettaen että maailmamme on *minimaalinen* P:n toteuttava maailma, ks. Chalmers 2004b, 280–2; 2009, 193):

- (1) $P \rightarrow Q$ on a posteriori.
 - (2) Jos $P \rightarrow Q$ on a posteriori, niin $P \rightarrow Q$ on 1-kontingenti.
 - (3) Jos $P \rightarrow Q$ on 1-kontingenti, niin $P \rightarrow Q$ on 2-kontingenti.
 - (4) Jos $P \rightarrow Q$ on 2-kontingenti, niin fysikalismi on väärässä.
-
- (5) Fysikalismi on väärässä.

(1) on edeltävä episteeminen väite. (2) on ilmeisesti totta niin kuin edellä todettiin. (4) on samoin ilmeisesti totta (ks. kuitenkin Piccinini 2017). Kuten kuviteltavuusargumentissakin, vakavin vastaväite kohdistuu kohtaan (3), jonka mukaan siitä, että $P \rightarrow Q$ on episteemisesti kontingenti seuraa, että $P \rightarrow Q$ on myös metafyyssisesti kontingenti. Väite on tosi vain väitteille, jotka ovat *semanttisesti neutraaleja* (Chalmers 2004b, 282), eli joiden episteemiset intensiot ovat samat kuin niiden subjunktiiviset intensiot.

Tässä Chalmersin päättelyä tukee intuitio siitä, että näennäinen ja todellinen yhtyvät

¹² Ks. luku 4.4 fenomenaalisisista arvostelmista.

¹³ Tämä on myös vasta-argumentti väitteeseen, jonka mukaan Marylla vain puuttuu tarvittavat fenomenaaliset käsitteet mutta ” $P \rightarrow Q$ ” on kyllä a priori, joten tietoargumentti kaatuu. (Muita Chalmersin vastineita tyypillisiin tietoargumentin vastaväitteisiin: ks. Chalmers, 2004, 284–293.)

fenomenaalisten ominaisuuksien tapauksessa: koska Q on puhdas fenomenaalinen käsite, sen episteeminen ja subjunktiivinen intensio ovat identtiset ja Q on semanttisesti neutraali. P kuitenkin ei ole semanttisesti neutraali, sillä jossain mahdollisessa maailmassa fysikaalisilla ominaisuuksilla voisi olla erilaisia intrinsisiä ominaisuuksia (subjunktiivisen intension poimima ilmiö) jotka kuitenkin toteuttavat samoja kausaalisia rooleja kuin omassa maailmassamme (episteemisen intension poimima ilmiö).

Jotta väitteen (3) materiaallinen implikaatio olisi epätosi,

(i) täytyy olla maailma w missä väitteen $P \rightarrow Q$ episteeminen intensio on epätosi ja sen subjunktiivinen intensio tosi.

Tällöin siis maailmassa w

(ii) Q:n episteeminen intensio on epätosi ja P:n episteeminen intensio on tosi, ja Q:n subjunktiivinen intensio on epätosi ja P:n subjunktiivinen intensio epätosi.

(Q:n episteeminen ja subjunktiivinen intensio ovat tietenkin samat.) Tällaisen maailman w täytyy olla kausaaliselta rakenteeltaan samanlainen kuin omamme, mutta sen intrinsiset ominaisuudet ovat erilaiset. Tämän eron vuoksi Q:n totuusarvo maailmassa w eroaa Q:n totuusarvosta omassa maailmassamme. Eli toisin sanoen, mikrofysikaalisen maailmamme kausaalinen rakenne ei ole syynä Q:n totuuteen, vaan maailmamme intrinsiset ominaisuudet – ts. saamme pan(proto)psykismin. (Chalmers 2004b, 283.)

Toisin sanoen, argumentti osoittaa Chalmersin mukaan, että joko fysikalismi on väärässä tai panpsykismi on totta – mikä sekään ei ole fysikalismia sen tavallisessa mielessä – ja näin argumentti on kummassakin tapauksessa vahva.

4 Chalmersin tietoisuusteoria

Koska kokemuksellinen tietoisuus on erityisasemassa, Chalmers esittää, jos edeltävät argumentit hyväksytään, kokemuksellisuus on otettava *fundamentaalisena*, jonakin joka tulee lisäyksenä ontologiaamme. Tietoisuus voi siis olla tällainen fysikaaliseen redusoitumaton fundamentaalinen ominaisuus, tai jotkin *protofenomenaaliset* ominaisuudet joista tietoisuus rakentuu, ja lisäksi täytyy olla näitä ominaisuuksia ohjaavia *fundamentaalisia lakeja* (Chalmers 1995, 210; 1996, 126; 2002a, 261). Jos Chalmersin vaikeaa ongelmaa ohjaavat intuitiot ovat pitäviä ja jos ensimmäisen persoonan tiedolla on episteeminen erityisasema joka tuottaa meille tietoa, joka ei ole osa fysikalistista tarinaa, Chalmersin 90-luvulla kehittämä tietoisuuden teoria näyttää olevan seurauksena. Esitän

sen lyhyesti seuraavaksi.

4.1 Metodologiset periaatteet

Luvun 3 argumenttien perusteella tietoisuus ei supervenoi loogisesti mutta kylläkin luonnollisesti fysikaalisen yllä. Toisin sanoen, on kehitettävä ei-reduktiivinen tietoisuuden teoria, johon kuuluu joitakin supervenienssilakeja, *psykofyysisiä lakeja*, jotka yhdistävät tietoisuuden ja fysikaaliset tosiasiat (Chalmers 1996, 213–5). Chalmersin mukaan tietoisuuden teorian tulee löytää yksinkertaisia, perustavia lakeja, hyväksyen samalla että jotakin jouduttaneen ottamaan *primitiivinä* (kuten Newtonin painovoimateoriassa painovoiman olemassaolo) ”eikä selityksellisestä yhteydestä tule aivan yhtä vahva kuin reduktiivisessa selittämisessä” (Chalmers 1997, 23).

Koska Chalmersin tietoisuuden teorian kulmakivenä on ensimmäisen persoonan tieto, emme voi testata psykofyysisiä hypoteeseja samaan tapaan kuin fysikaalisia hypoteeseja (Chalmers 1996, 215). Chalmersin mukaan saamme kuitenkin havaintoaineistoa omista kokemuksistamme, minkä lisäksi meillä on tietoa samanaikaisista fysikaalisista prosesseista, ja näiden säännönmukaisuuksien perusteella voimme abduktiivisesti koettaa löytää yksinkertaisimmat lait, jotka saavat aikaan nämä säännönmukaisuudet (Chalmers 1995, 211; 1996, 216). Myös normaalit tieteellisten teorioiden muodostusperiaatteet kuten yksinkertaisuus, koherenssi, uskottavuus, kauneus, lakien yleispätevyys jne. on huomioitava hyödyllisinä rajoitteina.

Tutkimuksessa on Chalmersin mukaan noudatettava seuraavaksi esitettäviä metodologisia periaatteita, joita ilman niitä tietoisuuden teorian muodostaminen ei olisi mahdollista (Chalmers 1996, 219).

Ihmisten raporttien todenmukaisuus

Kun toiset, kanssamme kognitiivisesti samankaltaiset ihmiset kertovat kokemuksistaan, pidämme uskottavana, että heidän kertomuksensa tosiaan ovat merkkejä tietoisista kokemuksista, ja heidän kokemuksia koskevat raporttinsa pääosin vastaavat heidän tietoisien kokemustensa sisältöä. Ei ole tällöin mitenkään epätieteellistä ottaa vakavasti tätä epistemiä tilaa (Levine 1994, 121). Tämä luottamus *intersubjektiivisen* (Velmans 2007, 717) tiedon mahdollisuuteen on ensisijaisesti metodologinen periaate: jos tämä ei pätsisi, Chalmers sanoo, mikä tahansa olisi mahdollista, maailma ei olisi järjellä käsitettävä paikka eikä tietoisuuden teorian muodostaminen olisi lainkaan mahdollista (Chalmers 1996, 216–7; 2004a, 1117).

Luotettavuusperiaate

Kysyessämme muilta heidän kokemuksistaan kuuntelemme heidän *fenomenaalisia arvostelmiaan* (kuten ”minulla on nyt punaisen aistimus”). Sovellettava metodologinen periaate tässä on *luotettavuusperiaate*, jonka mukaan nämä arvostelmat ovat pääosin paikkansa pitäviä (Chalmers 1996, 175–6, 218–9).

Havaittavuusperiaate

Kun meillä on kokemus, kykenemme nähtävästi myös muodostamaan arvostelman siitä. Vaikka emme muodostaisi arvostelmaa, meillä on kuitenkin *kyky* havaita kokemus: on vaikea kuvitella kokemusta joka meillä olisi, mutta jota emme periaatteessakaan voisi havaita. Tätä tutkimusta ohjaavaa periaatetta Chalmers kutsuu *havaittavuusperiaatteeksi* (Chalmers 1996, 219).

4.2 Psykofyysiset periaatteet ja lait

Rakenteellisen koherenssin periaate

Psykologisen tietoisuuden funktioihin liittyy tietoisuus jonkin läsnäolosta, tai täsmennyttymmin: suora globaali käytettävyyys (Chalmers 1995, 212), *a-tietoisuus*¹⁴. Sen sijaan *fenomenaaliseen tietoisuuteen* liittyy tietoisuuden kokemus. On jonkinlaista olla minä, siihen liittyy subjektiivinen näkökulma, jota ei voi kuvata psykologisen tietoisuuden funktioilla. Jatkossa ”tietoisuudella” tarkoitan aina fenomenaalista tietoisuutta Chalmersin tarkoittamassa merkityksessä, jollen nimenomaan toisin mainitse.

Chalmers erotteli aluksi (Chalmers 1996, 218–225) toisistaan 1) tietoisuuden ja a-tietoisuuden koherenssin, sekä 2) rakenteellisen koherenssin periaatteet, mutta alkoi myöhemmin puhua näistä yhteisnimikkeellä *rakenteellisen koherenssin periaate* (esim. Chalmers 1997, 30).

Chalmers esittää, että aina kun on tietoisuutta jostakin, on myös siihen liittyvä a-tietoisuus¹⁵. Tämä a-tietoisuus on tietoisuuteen liittyvä psykologinen ominaisuus, tila jossa jokin informaatio on suoraan ja globaalisti käytettävissä käyttäytymisen ohjaamiseen ja verbaalisten raporttien muodostamiseen (Chalmers 1995, 212 ja 225; 1996, 220). Tämä

¹⁴ *Awareness* (Chalmers 1996, 28). Chalmersin käsite ei vastaa sanan normaalia merkitystä englannin kielessä, eikä sille ole myöskään mitään kovin luontevaa suomennosta. *Tajunta* olisi yksi mahdollisuus, mutta sekin kantaa mukanaan kokemuksellisuuteen liittyviä sivumerkityksiä.

¹⁵ Samoin Blockin mukaan pääsy-tietoisuus ja fenomenaalinen tietoisuus ovat lähes aina läsnä tai poissa yhtä aikaa, ja siksi ne tavallisesti virheellisesti niputetaankin yhteen sekakäsitteeksi (Block 1995, 242).

määritelmä on hyvin lähellä sitä, mitä Block (1995) nimittää pääsy tietoisuudeksi.

Esimerkiksi visuaaliseen kokemukseeni kirjasta pöydälläni liittyy funktionaalinen *havainto* kirjasta; samoin jokainen yksityiskohta kokemuksesta on samoin edustettuna a-tietoisuudessa, mistä todistaa se, että voin muodostaa arvostelmia jokaisesta niistä. Kohteen ei tarvitse olla todellinen: esimerkiksi hallusinaatiossa minulla on sekä visuaalinen havainto että sitä koskeva arvostelma.

Chalmers määrittelee lisäksi a-tietoisuuden käsittämään vain *läsnä olevat ajatukset* (occurrent thoughts) – siihen ei siis kuulu esimerkiksi tietoni Suomen nykyisestä presidentistä, joka täytyy ”erikseen kutsua tietoisuuteen”. Näin Chalmers saa rakennettua symmetrian, jossa a-tietoisuuteen vastaavasti liittyy aina vastaava tietoinen kokemus (Chalmers 1996, 221–2).

Chalmersin mukaan ensimmäisen ja kolmannen persoonan tietoisuutta koskevien havaintojen välillä on selvä korrelaatio. Esimerkiksi näkökenttä on rikkaasti strukturoitu: minulla ei ole vain kokemuksia väriläiskistä, minulla on kokemus muodoista, suhteista ja etäisyyksistä, ja lisäksi koetuilla väreillä on intensiteetin, sävyn ja saturaation vaihtelua niin että kun aistin jonkin värin, läsnä on implisiittisesti myös kaikki ne värit joita en sillä hetkellä näe. Tietoisessa kokemuksessamme on *erottelurakenne*, ja tämä erottelurakenne on vastaavasti edustettuna kognitiivisesti, paitsi hermostollisessa tila-avaruudessa (P. M. Churchland 1989, 102), myös a-tietoisuuden rakenteessa (Chalmers 1995, 212; 1996, 223). Tästä todistaa taaskin se, että voin muodostaa arvostelmia näistä eroista ja että tämä informaatio on käytettävissä ohjaamaan käyttäytymistäni. Samankaltainen erottelustrukturi esiintyy äänissä, mauissa jne. Lisäksi, vaikka kvalitatiiviset kokemukset (kvaliat) olisivatkin sanoinkuvaamattomia, niiden väliset relaatiot kuten käsitys kokemuksen intensiteetistä, kokemuksen resoluutiosta jne. eivät ole. Ja nämä relaatiot vastaavat relaatioita a-tietoisuudessa.

Edeltävän a-tietoisuuden määritelmän mukaisesti jokaiseen a-tietoiseen tilaan liittyy aina vastaava kokemuksellinen tila, joten ”vaikuttaa uskottavalta” (Chalmers 1996, 225), että koherenssi toimii myös toiseen suuntaan. Näin saadaan symmetrinen rakenteellisen koherenssin periaate.

Chalmersin mukaan rakenteellisen koherenssin periaate toimii metodologisena perustana tietoisien kokemuksen empiiriselle tutkimukselle, niin että esimerkiksi kognitiivisen väriavaruuden rakenteen tutkiminen antaa suoraan tietoa kokemuksellisen väriavaruuden

rakenteesta. Rakenteellisen koherenssin periaate ei kuitenkaan silloita selityksellistä kuilua. Kuten aikaisemminkin, mikään rakenteellinen kuvaus a-tietoisuudesta ei selitä, miksi siihen liittyy myös tietoisuutta. Periaate yksinkertaisesti *olettaa* tietoisuuden olemassa olon, ei selitä sitä. (Chalmers 1996, 233–5.)

Chalmersin mukaan aina kun empiirisistä tutkimustuloksista tehdään päätelmiä tietoisuudesta, jokin esiempiirinen, *silloittava periaate* tekee kaiken todellisen työn, toimii *episteemisenä vipuvartena*. Tämä episteeminen vipuvarsi ei itse ole empiirisesti testattavissa kolmannen persoonan näkökulmasta, se vain toimii taustaolettamuksena. (Chalmers 1996, 236; 1998, 220.) Chalmersilla tällaisena episteemisenä vipuvartena toimii rakenteellisen koherenssin periaate. Monissa tapauksissa näitä silloittavia periaatteita voi pitää neutraalisti vain korrelaatioperiaatteina, ottamatta kantaa niiden kausaaliseen tai ontologiseen merkitykseen (Chalmers 2004a, 1116).

Koska, kuten edelle on esitetty, tietoisuus näyttää liittyvän tiiviisti ja *lainomaisesti* a-tietoisuuteen omassa tapauksessani ja edellä esitettyjen päättelyiden perusteella myös muilla ihmisillä, Chalmers katsoo perustelluksi tehdä *hypoteesin*: tietoisuuden ja a-tietoisuuden (rakenteellinen) koherenssi *on luonnonlaki*, jota pätee yleisesti (Chalmers 1996, 242). Viime kädessä Chalmersin perusteluna lain postuloimiselle toimii ennen kaikkea havainnot korrelaatiosta ensimmäisen persoonan tapauksessa: kyse ei voi olla vain yhteensattumista, Chalmers kirjoittaa (Chalmers 1996, 243).

Organisatorisen invarianssin periaate

Chalmers väittää, että tietoisuus nousee a-tietoisuuden *hienojakoisesta funktionaalista organisaatiosta* (fine-grained functional organization, FFO), joka on järjestelmän kausaalisen vuorovaikutuksen abstrakti muoto; ”hienojakoinen” tarkoittaa tarkastelutasoa, jolla (esimerkiksi) aivojen vaikutus käyttäytymiseen näkyy (Chalmers 1995, 213; 1996, 248). Näin saadaan Chalmersin dualistinen, ei-reduktiivinen funktionalismi.

Chalmers arvelee, että tämä ei kuitenkaan ole *fundamentaalin* luonnonlaki, sillä fundamentaalilaeissa ei yleensä ole tekijöinä niin epämääräisesti määriteltyjä käsitteitä kuin a-tietoisuus. Siirryttäessä kauemmaksi inhimillisestä kognitiosta tämä laki alkaa asteittain pettää, juuri a-tietoisuuden määritelmän epämääräisyyden takia (Chalmers 1996, 246). Myös vastaava Blockin pääsy tietoisuus on ryväs-käsite (Block 1995, 231).

Erilaiset järjestelmät voivat realisoida saman funktionaalisen organisaation. Kun kaksi järjestelmää realisoivat saman FFO:n ja ovat samassa tilassa tarkasteluhetkellä, ne ovat

toistensa funktionaaliset isomorfit. (Chalmers 1996, 247–8.)

Edeltävästä seuraa *organisatorisen invarianssin periaate*: jos järjestelmällä on tietoisia kokemuksia, toisella järjestelmällä, jolla on sama FFO, on myös kvalitatiivisesti identtiset tietoiset kokemukset. Tietoisuus on organisatorinen invariantti. Jos¹⁶ Blockin esittelemän ajatuskokeen Kiinan kansa on kausaaliselta organisaatioltaan identtinen tietoisien olennon kanssa, Kiinan kansasta muodostettu kokonaisuudella on tietoisia kokemuksia. Samoin tietokoneella, johon on kopioitu tietoisien olennon hienojakoinen kausaalinen organisaatio. Tietoisuus syntyy funktionaalisesta organisaatiosta – mutta ei ole funktionaalinen tila kuten reduktiivisessa funktionalismissa.

Todistellakseen organisatorisen invarianssin periaatetta Chalmers esittää i) *hiipuvien kvalioiden argumentin* ja ii) *tanssivien kvalioiden argumentin* (Chalmers 1995, 214–5; 1996, 249–74). Näiden tarkoituksena on osoittaa, että puuttuvat ja vaihdetut kvaliat eivät ole mahdollisia *luonnollisesti*. Argumentit eivät toimi reduktiivisen funktionalismin puolustukseksi koska 1) hiipuvat/tanssivat kvaliat näyttävät olevan kuitenkin loogisesti mahdollisia, ja koska 2) niihin sisältyy *empiirinen* premissi, jonka mukaan tietoista kokemusta on (Chalmers 1996, 274–5).

Chalmersin realismi on substanssimonismia, mutta *ominaisuusdualismia*. Tietoisuus supervenoi luonnollisella (mutta ei loogisella) välttämättömyydellä fyysikaalisen yllä: yksilön tietoisessa kokemuksessa on ominaisuuksia jotka riippuvat lainomaisesti yksilön fyysikaalisista ominaisuuksista, joidenkin kontingenttien luonnonlakien ansiosta (Chalmers 1996, 125). Chalmers edustaa *ei-reduktiivista funktionalismia*, jonka mukaan funktionaalinen organisaatio riittää määrittelemään tietoisien kokemuksen *luonnollisella* välttämättömyydellä, vaikkakaan tietoisuutta ei voi redusoida funktionaaliseen organisaatioon (Chalmers 1996, 275).

Rakenteellisen koherenssin ja organisatorisen invarianssin periaatteet eivät selitä esimerkiksi *minkälaisesta* organisaatiosta tietoisuus voi nousta, joten ne eivät kelpaa tietoisuuden teorian fundamentaalilajeiksi. Tällaiseksi laiksi Chalmers esitti 90-luvulla hyvin spekulatiivisen *informaation kaksoisaspektiteorian*, jonka mukaan (Shannonin klassisen määritelmän mukainen) informaatio on kaikkialla läsnä olevaa ja jotkin informaatiotilat realisoituvat isomorfisesti sekä fenomenalisesti että kokemuksen

¹⁶ Tietenkään meidän maailmassamme edes Kiinan kansan väkiluku ei riittäisi toteuttamaan ihmisaivojen hienojakoista funktionaalista organisaatiota ajatuskokeen kuvaamalla tavalla.

fysikaalisessa perustassa (Chalmers 1995, 216; 1996, 278–87). Tämä myös johti Chalmersin kiinnostumaan *panprotopsykismistä*, jonka mukaan eriasteista tietoisuutta on kaikkialla. Chalmers näyttää myöhemmin menettäneen kiinnostuksensa informaation kaksoisaspektiteoriaan, vaikka näkeekin yhä esimerkiksi Giulio Tononin informaation integraatioteorian (ks. esim. Tononi & Koch 2015) potentiaalisesti selitysvoimaisena.

4.3 Epifenomenalismi

Mikäli fenomenaalinen tietoisuus on jotakin *lisää* fysikaaliseen todellisuuteen ja hyväksytään fysikaalisen kausaalinen sulkeuma, joudutaan joko hyväksymään kausaalinen ylideterminaatio, tai hyväksyttävä, että fenomenaalisen tietoisuuden ominaisuuksilla ei ole mitään kausaalista vaikutusta fysikaaliseen. Chalmersin rakennelma ei salli samanlaista pääsyä tietoisuuden kausaaliseen voimaan kuin Searlen biologinen naturalismi, jonka mukaan tietoisuuden kvalitatiiviset, ensimmäisen persoonan näkökulmaan liittyvät prosessit ovat nekin biologisia (Searle 2004, 113).

Fysikaalisen sulkeumaa ei vakavissaan ehdota juuri kukaan hylättäväksi, sillä sen hylkäämällä voisi postuloida melkein mitä vain ei-empiirisesti mitattavia olioita fysikaalisten tapahtumien syiksi. Kausaalinen ylideterminaatio taas vaikuttaa mielivaltaiselta: miksi olettaa useampia riittäviä syitä, jos selitykseen riittää yksi? Chalmers myös hylkää molemmat edelliset vaihtoehdot (Chalmers 1996, 151–2). Epifenomenalismi taas antaa vaikutelman, että kokemuksellinen tietoisuus ja sen sisällöt jotenkin vain ”roikkuvat mukana” ilman minkäänlaista kausaalista voimaa, ja herättää sekin kysymyksen, miksi epifenomenaalisia ilmiöitä lainkaan olisi olemassa. Tämä näyttää vahvalta argumentilta Chalmersin tietoisuusteoriaa vastaan, ja sitä on myös usein käytetty (esim. Dennett 1991, 405; Frankish 2016a, 25; Revonsuo 2010, 10; Yudkowsky 2008).

Epifenomenalismia vastaan esitetään usein evolutionaarinen argumentti: miksi millekään järjestelmälle kehittyisi epifenomenaalisia ominaisuuksia, jos niillä ei ole eikä voi olla minkäänlaista vaikutusta yksilön tai lajin kehittymiseen? Chalmersin mukaan tämä ei kuitenkaan muodosta ylimääräistä ongelmaa, sillä evoluution prosessit ovat nekin fysikaalisia prosesseja ja tietoisuuden erityisasema säilyy samana kuin aiemminkin: miksi nämä fysikaaliset prosessit synnyttävät tietoisuuden? Tässä Chalmers päättelee samaan tapaan kuin Jackson (1982, 134), jonka epifenomenalisin puolustus yhtyy monissa kohdin Chalmersin näkemyksiin: jos psykofysikaaliset lait määräävät fenomenaalisten ominaisuuksien ilmenemisen fysikaalisten mukana, tietoisuus seuraa fysikaalisen

kehityksen mukana sivutuotteena¹⁷ (Chalmers 1996, 120; 2002a, 264).

Vaikuttaisi myös, että epifenomenaaliset kvaliat eivät voi olla toisten ihmisten käyttäytymisen syynä, joten emme voisi päätellä kvalioiden olemassaoloa muilla ihmisillä. Jacksonin vastaväite tähän on verrata lehti uutiseen: jos hän lukee *The Times*ista Spursin voittaneen ottelun, hän voi perustellusti päätellä, että *The Telegraph*issa on myös artikkeli, joka kertoo Spursin voittaneen, vaikka *The Telegraph*in artikkelin syynä ei olekaan *The Times*in artikkeli (Jackson 1982, 134).

Chalmers spekuloi, että kausaliteetin syvempi käsittäminen saattaisi antaa perusteita epifenomenalismiin torjumiseen (Chalmers 1996, 151). Kausaliteetti ei supervenioi loogisesti fysikaalisten tosiasioiden yllä, kuten jo Humen kritiikki osoittaa, eivätkä myöskään fysiikan lait (Chalmers 1996, 86)¹⁸ Eräänä ratkaisuna Chalmers pohtii ehdotusta, jonka mukaan fenomenaaliset ominaisuudet realisoivat kausaliteetin. Tällä tavalla kokemuksella olisi tärkeä rooli kausaatioissa, joskin tämä johtaisi panpsykismiin ja sen kombinaatio-ongelmiin. Samaan johtaa Chalmersin myöskin esittämä (ja suosima) ajatus kokemuksellisuudesta fysikaalisen maailman intrinsisenä piirteenä, joka ajatellaan lisäksi fysiikan lakien kuvaamiin struktuureihin.

Chalmers toisaalta ei halua kuvailla näkemystään epifenomenalistiseksi mutta samalla myöntää epifenomenalismiin ominaisuusdualismiin väistämättömäksi ja kiusalliseksi seuraukseksi (Chalmers 1996, 159–60). Siinä ei ole mitään varsinaisesti ristiriitaista, se vain on vahvasti intuition vastaista. Mutta epäintuitiivisuus ei ole peruste sen hylkäämiseen, tarvitaan riippumaton syy, jota ei näytä löytyvän (Chalmers 1996, 159).

Mutta fenomenaaliset arvostelmat näyttävät asettavan erityisen vaikean ongelman epifenomenalismille: jos tietoisuudella ei ole lainkaan kausaalista voimaa, miten meillä voi olla tosia uskomuksia siitä? Chalmersin näkemyksen mukaan ei kuitenkaan tarvitse olettaa kausaalista yhteyttä tietoisuuden ja sitä koskevien arvostelmien välille, vaan *vahvempi* yhteys, *konstituutio*. (Chalmers 2002a, 2003.) Tätä käsitellään seuraavassa luvussa.

¹⁷ Mørchin (2018) mielestä tällaiset yksisuuntaiset psykofyysiset lait eivät ole sen selittämättömpiä tai epätodennäköisempiä kuin interaktionistisen selityksen vaatimat kaksisuuntaiset lait tai fysikalismiin vaatimat identiteetit.

¹⁸ Erilaisista yrityksistä selvittää kausaliteetin ongelmaa ks. Price ja Corry 2007.

4.4 Fenomenaaliset arvostelmat

Intuitiot, joille Chalmersin tietoisuusteoria perustuu, näyttävät Chalmersin argumentoinnin perusteella jättävän fenomenaaliset arvostelmat merkilliseen asemaan. Nehän koskevat fenomenaalisen tietoisuuden tiloja, mutta kuuluvat kuitenkin psykologisen tietoisuuden ilmiöihin. Ne voidaan siis selittää funktionaalisesti ja reduktiivisesti, joten nähtävästi zombeilla voisi olla samanlaisia uskomuksia omista tietoisuuden tiloistaan ja ne kykenisivät raportoimaan niistä identtisesti meidän kanssamme (”Mitä? Tietenkin minä olen tietoinen olento!”), vaikka niiltä puuttuu täysin kokemuksellinen tietoisuus. Tämän paradoksin Chalmers koettaa selittää pois ensin *The Conscious Mindissa* ja sitten uudelleen rakennetulla fenomenaalisten arvostelmien teoriallaan (Chalmers 2003).

Chalmersin alkuperäisen jaottelun mukaan fenomenaalisia arvostelmia on kolmea tyyppiä. *1. tason arvostelmat* koskevat eivät kokemusta itseään vaan kokemuksen kohdetta. Kun näen punaisen läiskän seinässä, 1. tason arvostelma kuuluisi ”Näky jotakin punaista”. Ja a-tietoisuuden sisältö koostuu näiden 1. tason arvostelmien sisällöstä. *2. tason arvostelmat* koskevat tietoista kokemusta itseään. Edellisessä esimerkissä 2. tason arvostelma voisi olla ”Minulla on punaisen aistimus”. *3. tason arvostelmat* koskevat tietoista kokemusta *tyyppinä*. Tällaisia arvostelmia ovat esimerkiksi ”Punaisen kokemus on sanoinkuvaamaton”. (Chalmers 1996, 175–6.)

The Conscious Mindissa Chalmers vasta hahmotteli teoriaa, jonka mukaan fenomenaalisen käsitteen sisältöä ja vastaavaa uskomusta *konstituo*i osittain alla oleva fenomenaalinen ominaisuus (Chalmers 1996, 174 ja 204). Vuonna 2003 hän muotoili teorian tarkemmin ja nimesi sen *fenomenaaliseksi realismiksi*, jota esittelen seuraavassa.

Fenomenaalinen realismi

Fenomenaalisen realismin mukaan on todella olemassa fenomenaalisia ominaisuuksia, jotka eivät ole redusoitavissa käsitteellisesti fysikaalisiin tai funktionaalisiin ominaisuuksiin ja jotka tekevät mentaalista tiloista sellaisia, että on jonkinlaista olla niissä tiloissa. Lisäksi on olemassa totuuksia, jotka koskevat sitä millaista on olla subjekti ja jotka eivät seuraa a priori subjektin fysikaalisista tai funktionaalista totuuksista, ja ensimmäisen persoonan fenomenaaliset uskomukset ilmaisevat näitä totuuksia. (Chalmers 2003, 220–1; 2004b, 269.) Chalmersin mukaan fenomenaalinen realismi sisällyttää itseensä B-tyypin materialismin, mutta fenomenaalinen realismi ei ole yhteensopiva A-tyypin materialismin kanssa. (Chalmers 2003, 222; 2004b, 270.)

Relationaaliset, demonstratiiviset ja puhtaat fenomenaaliset käsitteet

Kun katson punaista omenaa, tämä kokemus instantioi fenomenaalisen kvaliteetin R, jota voimme kutsua fenomenaaliseksi punaisuudeksi; huomioidessani kokemustani ajattelen ”minulla on kokemus sellaisesta-ja-sellaisesta ominaisuudesta” (Chalmers 2003, 223). Jatkossa tässä luvussa ominaisuudet (esim. R) merkitään normaalilla fontilla ja käsitteet (esim. **R**) lihavoituina.

Voin viitata tähän ominaisuuteen julkisen kielen käsitteellä ”punainen”, tai julkisen kielen ilmaisulla ”fenomenaalinen punainen”. Näissä käsite on kiinnitetty relaatioon ulkoisessa maailmassa oleviin punaisiin esineisiin (kokemuksen aiheuttajina) ja relaatioon kokemuksiin (”punaisten” esineiden subjektissa tyypillisesti aiheuttamiin), niin kuin ne opitaan opittaessa julkisen kielen käsite ”punainen”. (Chalmers 2003, 224; 2004b, 271.)

Näitä ”fenomenaaliseen punaiseen” viittaavia relationaalisia fenomenaalisia käsitteitä on kaksi, joista ensimmäinen on *yhteisöllisesti relationaalinen käsite* **punainenC** (tyypillisten punaisten esineiden *yhteisön subjekteissa* yleensä aiheuttama fenomenaalinen ominaisuus), joka myös takaa käsitteen jaetun merkityksen yhteisössä. Lisäksi on *yksilöllinen relationaalinen käsite*, **punainenI** (tyypillisten punaisten esineiden *minussa* aiheuttama fenomenaalinen ominaisuus). Poikkeusyksilöitä (esim. mahdollisia väriavaruudeltaan käännettyjä subjekteja) lukuun ottamatta näiden referentit ovat samat.

Fenomenaaliset käsitteet voidaan poimia myös indeksikaalisesti: voin sanoa ”Tämä ominaisuus” tai ”tämän tyyppinen kokemus”. Nämä ilmaisevat demonstratiivisen (tai indeksikaalisen, ks. Chalmers 2003, 228) käsitteen **E** (”tämä ominaisuus, mikä se sitten sattuukaan olemaan”). (Chalmers 2003, 225; 2004b, 272.)

Käsitteet **punainenC**, **punainenI** ja **E** voivat kaikki viitata samaan fenomenaaliseen ominaisuuteen, ja jokaisessa tapauksessa relationaalisesti. Mutta lisäksi on eräs fenomenaalinen käsite, joka poimii punaisuuden suoraan, sen intrinsisen fenomenaalisen luonteen avulla. Tätä Chalmers kutsuu *puhtaaksi fenomenaaliseksi käsitteeksi*¹⁹ (pure phenomenal concept) (Chalmers 2003, 225; 2004b, 272.). Kun Mary pääsee huoneesta, hän oppii, että punaisen näkemiseen tyypillisesti liittyy kokemus *tällaisesta* ominaisuudesta. Merkitään tätä Maryn *tällaisuuden* käsitettä kirjaimella **R**.

¹⁹ Vastaa Levinen *fenomenaalisesti konstituoitua* ajatusta, erotuksena *fenomenaalisuuden saattamasta* ajatuksesta (Levine 2001, 158).

R on kvalitatiivinen käsite, jonka Mary oppii nähdessään ensi kerran punaista; se on sidottu a priori tietynlaiseen ominaisuuteen, eikä se ole sama kuin edellä esitetyt käsitteet. Tämä käy selväksi, kun huomataan, että hän myös saa uskomukset

punainenC = **R** (punaisten esineiden Maryn yhteisössä yleensä aiheuttama ominaisuus on *tällainen*)

punainenI = **R** (punaisten esineiden minussa aiheuttama fenomenaalinen ominaisuus on *tällainen*)

E = **R** (tämän tyyppinen ominaisuus on siis *tällainen*).

Keskeistä Chalmersin mukaan on, että yhtäläisyys **E** = **R** on kognitiivisesti merkittävä, se ei ole a priori tiedettävissä. Se yhdistää demonstratiivisen käsitteen kvalitatiiviseen käsitteeseen. Käsitettä **R** on vaikea ilmaista kielessä, sillä tavalliset ilmaisut kuten ”fenomenaalinen punaisuus” tai ”tämä kokemus” ilmaisevat aina jonkin muun käsitteen (tässä **punainenC** ja **E**, vastaavasti) (Chalmers 2003, 227).

Chalmersin kaksiulotteisessa semantiikassa käsitteillä **punainenC**, **punainenI**, **E** ja **R** on erilaiset *episteemiset* intensiot mutta sama *subjunktiivinen* intensio: niiden jokaisen subjunktiivinen intensio poimii fenomenaalisen punaisuuden kaikissa maailmoissa. Sen sijaan käsitteen **R** episteeminen intensio poimii fenomenaalisen punaisuuden kaikissa episteemisissä maailmoissa. (Chalmers 2004b, 274–5.)

Värikokemukseltaan invertoidun Maryn ajatuskoetta analysoidessaan Chalmers toteaa, että puhtaan fenomenaalisen käsitteen tapauksessa laadullinen kokemus konstituoi käsitteen ja vastaavan uskomuksen *episteemistä* sisältöä (Chalmers 2003, 229–33; 2004b, 276). Kuten Chalmers jo 1996 muotoili, tässä tapauksessa käsitteen referentti on ”sisällä käsitteen mielessä (sense)” (Chalmers 1996, 207; 1997, 22). Puhdas fenomenaalinen käsite on *episteemisesti jäykkä*, kun tavalliset jäykät käsitteet ovat vain subjunktiivisesti jäykkiä, eli ne poimivat saman referentin jokaisessa mahdollisessa maailmassa (käsitettynä kontrafaktuaalisena) (Chalmers 2003, 233).

Chalmers käyttää tätä puolustamaan Jacksonin tietoargumenttia. Tietoargumentin vastaväitteenä esitetään joskus, että fenomenaaliset käsitteet ovat indeksisiä (tai demonstratiivisia) käsitteitä, jolloin episteeminen kuilu on vain osa yleisempää kolmannen persoonan tiedon ja indeksisen (tai demonstratiivisen) tiedon välistä kuilua. Mutta **E** edellä on tämä indeksinen tai demonstratiivinen käsite, ja se ei ole sama kuin **R**, joten nämä tietoargumentin vastaväitteet eivät kestä. (Chalmers 2004b, 277–8.)

Fenomenaalisten uskomusten sisältö

Fenomenaalisen käsitteen sisältöä ja vastaavaa fenomenaalista uskomusta *konstituo* osittain alla oleva fenomenaalinen ominaisuus. Chalmers kutsuu tällaista fenomenaalista käsitettä ilmaisulla *suora fenomenaalinen käsite* (direct phenomenal concept) (Chalmers 2003, 235). Niin kuin puhtaiden fenomenaalisten käsitteiden tapauksessa yleensäkin, meillä ei ole julkista kieltä jolla voisi yksikäsitteisesti ilmaista suorien fenomenaalisten käsitteiden sisältöä. (Chalmers 2003, 240)

Kun Marylla on fenomenaalinen punaisen kokemus, hän muodostaa suoran fenomenaalisen käsitteen **R** ja sitten muodostaa uskomuksen *tämä kokemus on R*, osoittaen kyseistä fenomenaalista punaista kokemusta. Tällaista uskomusta Chalmers kutsuu *suoraksi fenomenaaliseksi uskomukseksi* (direct phenomenal belief) (Chalmers 2003, 236–8).

Puhtaita fenomenaalisia käsitteitä on muitakin. Esimerkiksi Marylle voi jäädä tietoa siitä millaista on nähdä tomaatteja, kun hän on jo palannut mustavalkoiseen huoneeseensa. Maryllä on nyt puhdas fenomenaalinen käsite, jota Chalmers nimittää *viipyväksi fenomenaaliseksi käsitteeksi S* (standing phenomenal concept) (Chalmers 2003, 239).

Näillä käsitteillä Mary voi muodostaa erilaisia fenomenaalisia uskomuksia: suoria fenomenaalisia uskomuksia kuten **E = R** (koska tuttuus ja konstituutio), apriorisia uskomuksia kuten **R = S** (a priori sisällön vuoksi), ja aposteriorisia fenomenaalisia uskomuksiin kuten **S = punainenC** ja **S = punainenI** (Chalmers 2003, 259)

Puhtaiden fenomenaalisten käsitteiden ja uskomusten sisältö ei redusoidu fysikaaliseen ja funktionaaliseen, koska kokemus konstituoii tätä sisältöä. Samoin eivät redusoidu uskomukset ja käsitteet, joissa nämä puhtaat fenomenaaliset käsitteet/uskomukset ovat osina. Tällainen redusoitumaton jäännös on siis hyvin laajalti käsitteistössä. (Chalmers 2003, 265–6)

Chalmersin esittää, että suoran fenomenaalisen uskomuksen määritelmästä seuraa

Erehtymättömyysteesi: suora fenomenaalinen uskomus ei voi olla väärässä.
(Chalmers 2003, 242.)

Tämä ei kuitenkaan koske kaikkia fenomenaalisia tiloja: esimerkiksi kun sovelletaan viipyvää puhdasta fenomenaalista käsitettä uuteen tilanteeseen, tai suoraa fenomenaalista käsitettä sovelletaan ominaisuuteen tai kokemukseen joka ei ole sama joka konstituoii sen (kun vaikkapa punaisen suora fenomenaalinen käsite on muodostettu näkökentän

vasemmalla puolella mutta sitä sovelletaan oikeaan puoleen. (Chalmers 2003, 242–4.) Voin myös erehtyä siitä, onko minulla suora fenomenaalinen uskomus kun esimerkiksi luulen viipyvää fenomenaalista uskomusta suoraksi fenomenaaliseksi uskomukseksi. Argumentointi ei siis tue väitettä, että *kaikki* fenomenaaliset tilat voisi tietää erehtymättömästi (Chalmers 2003, 244–5).

Miten sitten niistä puhtaista fenomenaalisisista käsitteistä, jotka erehtymättömyysteesi kattaa Chalmersin lukuisten varausten jälkeen, voi muodostaa erehtymättömiä? Chalmersin mukaan tämä perustuu epätavallisen läheiseen relaatioon, joka meillä on kokemuksemme fenomenaalisiin ominaisuuksiin. Hän kutsuu tätä relaatiota *tuttuudeksi* (acquaintance)²⁰, joka on perustavanlaatuisen episteeminen relaatio subjektin ja ominaisuuden instanssin välillä (Chalmers 2003, 248–50). Tähän nojautuen saadaan

Oikeutusteesi: Kun subjekti muodostaa suoran fenomenaalisen uskomuksen, joka perustuu fenomenaaliseen ominaisuuteen, tämä uskomus on prima facie oikeutettu siksi että ko. ominaisuus on tuttuusrelaatioissa subjektin. (Chalmers 2003, 249.)

Fenomenaalisten arvostelmien paradoksi

Fenomenaalisisella realismilla on paradoksaaliselta vaikuttava seuraus: sen mukaan kokemukset ovat kausaalisesti irrelevantteja fenomenaalisisille uskomuksille, mistä edelleen seuraa, että fenomenaalisiset uskomukset eivät voi olla tietoa. Chalmers toteaa tämä paradoksin hankalaksi, joka pitää jotenkin käsitellä. Ja näin hän tekee²¹.

Chalmersin mukaan paradoksiin johtaa oletus kausaalisesta yhteydestä kokemuksen ja fenomenaalisisen uskomuksen välille. Mutta vaikka sitouduttaisiin epifenomenalisiin, edellä esitetyn mukaan kokemuksen ja fenomenaalisisen uskomuksen välinen yhteys on *konstituutio*, joka on kausaalista vahvempi yhteys. Vaihtoehtoisesti epifenomenalisti voi vedota edellä esitettyyn tuttuusrelaatioon. (Chalmers 2003, 254–6.)

Chalmersin mukaan meidän tietoisuutta koskevat arvostelmamme voivat olla tosia, vaikka zombien täysin samanlaiset arvostelmat ovat virheellisiä: esimerkiksi Mary voi muodostaa uskomuksen **E = R**, mutta minkä uskomuksen zombi-Mary voisi muodostaa? Koska fenomenaalisisen ominaisuus konstituoi fenomenaalisisen käsitettä ja uskomusta, zombi-Maryn käsitettä **R** vastaava käsite on uskottavimman selityksen mukaan *tyhjä*:

²⁰ Tuttuus on fundamentalistisen tietoteorian keskeisiä periaatteita. Ks. myös Bertrand Russell (1910).

²¹ Yudkowsky Chalmersin haluttomuudesta kyseenalaistaa intuitiot, jotka ovat johtaneet tähän hankalaan tilanteeseen: “I would seriously nominate this as the largest bullet ever bitten in the history of time.” (Yudkowsky 2008.)

sillä ei ole sisältöä²². Joten zombi-Maryn uskomus ei voi olla tosi, koska sillä ei ole propositionaalista sisältöä. (Chalmers 2003, 257.)

Paradoksin ratkaisu näin ollen seisoo ja kaatuu sen varassa, voiko ensimmäisen persoonan tiedon episteemisen erityisaseman hyväksyä.

5 Realistien keskeiset intuitiot

Tutkielman tarkoituksena on käsitellä Chalmersin metodologiaan kohdistettua kritiikkiä ja varsinkin tapaa jolla hän (kuten muutkin realistit) tukeutuu intuitioihin. Ennen kritiikkiosioon siirtymistä aloitan luokittelemalla nämä intuitiot.

Kuten edeltävistä luvuista käy ilmi, Chalmers, kuten myös Nagel, Jackson ja Levine (osin myös Kripke) tukevat useasti argumentointiaan intuitiivisiin käsityksiin. He ovat toki hyvin tietoisia tästä ja yleensä perustelevat, että juuri nämä intuitiivisesti tosiksi havaitut ilmiöt ovat nimenomaan selitettävä asia: ilman niitä ei olisi mitään selitettävää, ei vaikeaa ongelmaa. Koko vaikea ongelma perustuu käsitykseen, jonka mukaan funktionaalisten selitysten jälkeen on lisäksi vielä jotakin muuta: subjektiivinen näkökulma, punaisen *tällainen punaisuus*, kokemus siitä *millaista on olla*, kokemuksellisuuden käsittäminen niin erillisenä materiaalisesta, että zombieargumentti tuntuu mahdolliselta.

Chalmersin totesi hänen kiistansa Dennettin kanssa koskevan eroja ensimmäisen persoonan fenomenologiaan liittyvissä perusintuitioissa (Chalmers 1997, 15). Samoin Dennett kirjoittaa:

To me one of the most fascinating bifurcations in the intellectual world today is between those to whom it is obvious – *obvious* – that a theory that leaves out the Subject is thereby disqualified as a theory of consciousness (in Chalmers' terms, it evades the Hard Problem), and those to whom it is just as obvious that any theory that *doesn't* leave out the Subject is disqualified. (Dennett 2001a, 229.)

Illusionistien intuitiot eroavat usein realistien intuitioista, mutta yhtä usein intuitio saattaa olla sama – illusionisti ei vain myönnä sille minkäänlaista perustavaa merkitystä. Tässä tutkielmassa käsittelyn alaisena ovat siis intuitiot, joiden varassa realistien argumentointi seisoo, ja niiden luokittelussa noudatan Chalmersin (2018a) jaottelua.

Ongelmaraporteiksi Chalmers kutsuu lausahduksia kuten ”On olemassa tietoisuuden vaikea ongelma”, ”Käyttäytymisen selittäminen ei selitä tietoisuutta”. Tietoisuuden

²² Vrt. taas Levine 2001, 159.

metaongelma käsittelee kysymystä, miten tällaiset ongelmaraportit syntyvät. Koska nämä ovat osa käyttäytymistä, *metaongelma* kuuluu Chalmersin teoriassa tietoisuuden helppoihin ongelmiin, vaikkakin Chalmers uskoo, että *metaongelman* ratkaisu voi auttaa myös vaikean ongelman ratkaisemisessa. (Chalmers 2018a, 7–8.) Samoin sen mikä selittää vaikean ongelman voi myös olettaa olevan keskeisessä osassa selitettäessä *metaongelmaa* (Chalmers 2018a, 36).

Chalmersin mukaan *metaongelma* ottaa puolueettoman kannan kysymykseen tietoisuuden olemassaolosta ja luonteesta (Chalmers 2018a, 10), mutta materialistinen *metaongelman* ratkaisuyritys on tyypillisesti jotakin jota Keith Frankish (2016a) kutsuu *illusionismiksi* ja jonka mukaan tietoisuus on jonkinlainen *illuusio*.

Verbaalisten raporttien sijaan varsinaiset kiinnostuksen kohteet ovat niiden perustana olevat tilat, ja näitä taipumuksia tehdä ongelmaraportteja Chalmers kutsuu *ongelma-intuitioiksi* (problem intuitions). Nämä voivat olla filosofisista argumenteista seuraavia arvioita mutta useimmiten *ongelma-intuitiot* ovat arvioita jotka heräävät *ennen* filosofisia argumentteja. *Ongelma-intuitioita* on ainakin seuraavat (Chalmers 2018a, 12–13), joista keskeisimpiä ovat seitsemän ensimmäistä.

1. *Metafyysiset intuitiot*, kuten dualistinen intuitio jonka mukaan tietoisuus on ei-fysikaalinen, tai fundamentalistinen intuitio jonka mukaan tietoisuus on fundamentaalinen. Esimerkki: ”tietoisuus on yllättävä ilmiö”.
2. *Selittävät intuitiot*, joiden mukaan tietoisuutta on vaikea selittää. Esimerkki: ”fysikaalisten prosessein ja tietoisuuden välillä on selityksellinen kuilu”.
3. *Tietointuitiot*, jotka koskevat ensimmäisen persoonan näkökulmasta saadun tiedon erityisasemaa. Esimerkki: ”ei ole mahdollista käsittää lepakon kokemuksellista tietoisuutta”.
4. *Modaali-intuitiot*, jotka koskevat sitä mikä on kuviteltavissa tai mahdollista. Esimerkki: ”zombit ovat kuviteltavissa”.
5. *Itseä koskevat intuitiot*, jotka koskevat itsen jatkuvuutta ja itsen ykseyttä.
6. *Kvaliteetti-intuitiot*, jotka koskevat erityisiä kokemukselle läsnä olevia kvaliteetteja.
7. *Esitysintuitiot* siitä miten em. kvaliteetit esittäytyvät meille.
8. *Arvo-intuitiot*, joiden mukaan tietoisuudella on erityinen arvo tai moraalinen merkitys.
9. *Kattavuusintuitiot*, koskien sitä millä järjestelmillä on tai voi olla tietoisuus.

Chalmersille metaongelmassa on kysymys siitä miten nämä ongelmaintuities voidaan selittää *aiheneutraalisti* (Smart 1959), sitoutumatta ominaisuusdualismiin tai materialismiin ja mainitsematta tietoisuutta tai sen kvalitatiivisia ominaisuuksia selityksessä. Tällaiset selitykset voivat olla vaikkapa representatiivisia, rationaalisia, historiallisia ja strukturaalisia, ja niitä voi yleensä myös testata empiirisesti. (Chalmers 2018a, 16.)

6 Tietoisuus illusiona

Ennen kuin siirryn Dennettin hyökkäykseen näitä ongelmaintuities kohtaan, esittelen yleisnäkymän realismin kohtaamaan vastarintaan. Tämän voi luokitella monella tavalla, esimerkiksi Chalmers on tehnyt useita jaotteluita, mutta tässä käytän Frankishin käsitettä *illusionismi*. Tämä kattaa useita eri näkemyksiä, joita yhdistää käsitys tietoisuuden tai ainakin sen joidenkin ongelmallisten piirteiden illusorisuudesta.

Dennettin mukaan illusionisti, sanan tavanmukaisessa merkityksessä, on erikoistunut kaikkeen mikä liittyy lavataikuuteen ja sen temppuihin. Lavataikuuden tapauksessa me oletamme, mikäli ei ole todistettu toisin, että ilmiö on selitettävissä jonkin vaikeasti toteutettavan tempun ja apuvälineiden avulla. Lavataikuuden selittämiseen ei riitä sanoa, että ”on vain kiistämättömän, intuitiivisen ilmeistä, että mikään mahdollinen tavallisten fyysikaalisten tapahtumien ketju ei voisi saada aikaan juuri nähtyä”. Samoin filosofit, jotka nojautuvat kiistämättömiin *intutioihin* tietoisuudestaan, eivät noudata hyvää tutkimusmetodia. Ensin on tutkittava vaihtoehdot. (Dennett 2016, 65–67.)

Frankishin mukaan tieteessä ei pidä olettaa mitään tarpeettoman radikaalia, mikä ei ole välttämätöntä ilmiön selittämiseksi, ja tämä on erityisesti muistettava, kun paine radikaaliin muutokseen tulee 1. persoonan epäluotettavasta näkökulmasta (Frankish 2016a, 24). Jos havaitsemme jotakin hyvin anomaalista, kuten telekinesiaa, joka ei tunnu sopivan tieteelliseen käsityksemme maailmasta, meillä on kolme vaihtoehtoa. 1) Voimme hyväksyä anomalian ja päätellä, että meidän tulee tehdä radikaaleja muutoksia tai laajennuksia teorioihimme; tämä oli Chalmersin ratkaisu *The Conscious Mindin* aikoihin. 2) Voimme hyväksyä kyseisen ilmiön olemassa olevaksi, mutta arvella, että se on aikanaan selitettävissä nykyisten teorioiden avulla. 3) Voimme väittää, että ilmiö on illusorinen, ja alkaa selvittää, miten tällainen illuusio syntyy. (Frankish 2016a, 12–13.)

Kaksi ensimmäistä näistä on *realistisia* näkökantoja, kolmas on *illusionistinen*.

Tietoisuuden tapauksessa anomalian asemassa näyttää olevan sen subjektiivinen,

fenomenaalinen aspekti, ja mahdolliset kolme suhtautumistapaa Frankishin mukaan ovat 1) *radikaali realismi*, 2) *konservatiivinen realismi* ja 3) *illusionismi* (Frankish 2016a, 13–14), jotka vastaavat lähes tarkalleen Chalmersin jaottelua 1) dualismiin²³, 2) B-tyypin materialismiin ja 3) A-tyypin materialismiin, vastaavasti.

Radikaali realismi sisältää Frankishin mukaan *ylimääräisiä* metafysisiä sitoumuksia. Radikaalin realistisen tieteen tarjoama dataa ei voi myöskään vertailla eri aikoina tai tarkistaa, että aiempia tapahtumia koskevat uskomuksemme ovat todenmukaisia, mikä tekee mahdottomaksi rakentaa radikaalin realismin tietoisuuden teoriaa. Russellin monismi taas vain lisäisi uskomuksen, jolla ei ole mitään muuta tehtävää kuin tukea uskomustamme, että meillä on suora pääsy fenomenaalisiin ominaisuuksiin. (Frankish 2016b, 286.)

Konservatiivisten realistien mielestä fenomenaaliset ominaisuudet tosiaan ovat kvalitatiivisia, mutta vain harhaanjohtavasti esitettyjä. Heidän tavallinen taktiikkansa on *fenomenaalisten käsitteiden strategia* (ks. esim. Chalmers 2006, 2018a), jonka lienee ensimmäisenä esittänyt Brian Loar (1990), ja joka tältä osin jakaa B-tyypin materialismin ongelmat (Frankish 2016a, 15–16, 25–26). Tätä ei tämän tutkielman puitteissa ole mahdollista käsitellä.

Illusionismi vaikuttaa radikaalia realismia taloudellisemmalta, elegantimmalta ja selitysvoimaisemmalta (Frankish 2016b, 285). Frankishin argumentti illusionismin puolesta on seuraava (Frankish 2016a, 27–28). Jos ominaisuus vastustaa kaikkia fysikaalisia selitysyrityksiä tai se voidaan havaita vain jostain tietystä perspektiivistä, yksinkertaisin selitys on, että se on illuusio. Niinpä, jos on vähäinkin epäily siitä, että olemme erehtyneet fenomenaalisen tietoisuuden olemassaolosta, on vahvat syyt tehdä abduktiivinen päätelmä, että tietoisuus on illuusio. Ja on hyviä syitä uskoa, että olemme erehtyneet, sillä meillä ei ole mitään introspektiivista keinoa tarkista introspektiomme paikkansapitävyyttä. Fenomenaalinen tietoisuus voisi siis olla illuusio, joten, abduktiivisesti, se on illuusio.

Vahva illusionismi väittää, että tietoisuus on illusorinen: olemme introspektiivisesti tietoisia sisäisistä kokemuksellisista tiloistamme, mutta tämä tieto on puolueellista ja vääristynyttä. Frankishin mielestä kyse ei silti ole eliminativismista, toisin kuin

²³ Mukaan lukien ”neutraali monismi, mystisismi ja vetoamiset uuteen fysiikkaan” (Frankish 2016a, 13).

esimerkiksi Katalin Balog (2016) väittää, sillä vahvakaan illusionismi ei kiistä kokemuksellisuuden olemassaoloa vaan väittää ainoastaan, että kokemuksellisuuteen liittyy vain *kvasi-fenomenaalisia* ominaisuuksia (Frankish 2016a, 21; 2016b, 263). Fenomenologiset käsitteet poimivat kyllä esiin kausaalisesti vaikuttavia ominaisuuksia, illusionisti vain kiistää arkipsykologian mukaisen mielipiteen, jonka mukaan ne ovat kvalitatiivisia (Frankish 2016a, 27).

Frankish väittää, että zombeilla on samanlainen kokemus siitä ”millaista on olla” kuin meillä, koska niillä on kanssamme samanlainen introspektiivinen mekanismi (Frankish 2016a, 23). Kuten myöhemmin nähdään²⁴, Dennettin näkemys on samankaltainen.

Heikon illusionismin mukaan jotkin fenomenaaliset ominaisuudet ovat illusorisia, mutta aitoja fenomenaalisia tiloja on kuitenkin olemassa. Eräs muoto heikosta illusionismista on Susan Blackmoren *delusionismi*. Blackmore kysyy, millaista ”on olla” niinä hetkinä kun emme kysy itseltämme millaista on olla, ja väittää, että meillä ei ole ”jatkuvaa tietoisuuden virtaa” – vain tällaisten kysymysten hetkillä konstruoituja tilapäisiä ajatuksista ja havainnoista koostuvia rakennelmia, jotka esittävät itsen jatkuvana tarkkailijana (Blackmore 2016). Frankishin mielestä heikko illusionismi tasapainoilee vaikeassa väli-maastossa ja on vaarassa romahtaa vahvaan illusionismiin – tai sitten sen on omaksuttava käsitys fenomenaalisuudesta jonakin vahvempana kuin kvasi-fenomenaalisuus. (Frankish 2016a, 15–16, 26). Myös Chalmersin mielestä illusionistin on oltava vahva illusionisti (Chalmers 2018a, 49–52).

Illusionistit esittävät erilaisia analogioita valottaakseen näkemystään. Georges Rey (2016) esittää, että vakiintuneet suhteemme maailmaan johtavat meidät projisoimaan vastaavia ominaisuuksia maailmaan ja introspektiossa itseemme. Dennettin (esim. 1991) mukaan tietoisuus on käyttäjäilluusio, hieman kuten tietokoneen käyttöliittymä. Yhteistä näille analogioille on, että niiden mukaan introspektio esittää meille vain vajavaisen, vääristyneen kuvan kokemuksistamme, jossa se esittää monimutkaiset fyysiset prosessit yksinkertaisina, fenomenaalisina ominaisuuksina (Frankish 2016a, 18). Aistimuksemme esittävät väärin reaktiomme, jotka koostuvat monimutkaisesta hermos-tollisesta toiminnasta, yksinkertaisena fenomenaalisena tuntemuksena. Tämän ei tarvitse merkitä *vikaa* aistimuksessa: vääristymä saattaa olla hyödyllinen tarvittavan vaikutuksen saamiseksi (pino sosiologisia tutkimuksia vanhempi-lapsi-suhteesta ei tuota samaa

²⁴ Ks. luku 8.4.1.

vaikutusta kuin *Kuningas Learin* näkeminen). (Frankish 2016b, 273–4.)

Illusionismia tukee myös *debunkkausargumentti* (Chalmers 1996, 186–7; 2018a, 44–45): jos on olemassa pääosin reduktiivinen, tietoisuudesta itsestään riippumaton selitys tietoisuutta koskeviin uskomuksiimme, ei-reduktionistiset uskomuksemme tietoisuudesta ovat perusteettomia. Chalmers uskoo kuitenkin voivansa puolustautua fenomenaalisten arvoistelmien teoriallaan: jos tietoisuus *konstituoii* fenomenaalisia uskomuksia, näiden uskomusten selitys ei ole riippumaton tietoisuudesta eikä debunkkausargumentti toimi – ja samoin jos hyväksytään, että välitön *tuttuutemme* tietoisuutemme kanssa oikeuttaa tietoisuutta koskevat uskomuksemme, fenomenaaliset uskomukset voivat pysyä oikeutettuina (Chalmers 2018a, 46).

Chalmersin mielestä pystymme silti selittämään ”epämukavan paljon” (Chalmers 2018a, 47) fenomenaalista uskomuksistamme vetoamalla ainakaan tietoisuuteen: jos metaongelmaprosessit *olisivat* voineet ilmetä ilman tietoisuutta, näyttää vain hyvältä onnelta, että niin ei ole käynyt. Samoin näyttää, että Chalmersin psykofyysiset lait ovat voimassa pelkästään hyvällä onnella – ainoa mikä estää meitä elämästä zombimaailmassa (Chalmers 2018a, 47–48). Chalmers pohtii pelastukseksi jonkinlaista *realisationismia*, jonka mukaan *vain* tietoisuus voi realisoida metaongelmaprosessit (Chalmers 2018a, 42 ja 49), ehkä Mørchin (2018) esittämällä tavalla: fenomenaalisten tilojen luonteeseen itseensä sisältyvät niiden kausaaliset voimat.

Chalmersin mielestä tämä ratkaisu sopisi sekä reduktionistille että ei-reduktionistille – mutta illusionismi tarjoaisi silti selvästi yksinkertaisemman selityksen, ilman yritystä kantaa ”fenomenaalisuuden” painolastia mukana.

Argumentteja illusionismia vastaan

Realistit sanovat, että illusionisti kieltää ilmeisen, sen mikä on explanandum (esim. Chalmers 2002a, 251). Levine esimerkiksi kysyy, kuka ottaisi vakavasti tietoisuuden teorian, joka kieltäisi lähtökohtaisesti vaikkapa ihmisen kielellisen käyttäytymisen olemassaolon (Levine 1994, 114)? Vertaus psykologisen tietoisuuden piirteisiin ei ole kovin toimiva, mutta joka tapauksessa illusionistit eivät näe kummassakaan tapauksessa ongelmaa: illusionisteille introspektion selostukset ovat vain aineistoa, joka vaatii tulkintaa ja arviointia, ja paras selitys sille voi olla tulkinta, joka kieltää sen luotettavuuden (Dennett 2003; Frankish 2016a, 30).

Toinen keskeinen illusionismin vastainen argumentti on, että on epä johdonmukaista

epäillä, että kokemuksemme ovat sitä miltä ne näyttävät, koska kvalitatiivisten tilojen tapauksessa ei ole eroa sillä miltä ominaisuus näyttää ja mitä se todella on. Illusionisti myöntää, että meillä ei ole keinoa korjata vilpittömiä kokemusraporttejamme ja meidän on annettava niille tässä mielessä auktoriteetin asema – mutta tämä auktoriteetti ei ole *episteeminen*, sen sijaan pikemminkin samanlainen auktoriteetti kuin mitä tarinanker-tojalla on tarinoihinsa (Dennett 1991, 81; Frankish 2016a, 33).

Kolmas tavallinen vastaväite liittyy illusion kokijaan. Jos siis ei ole olemassa kokevaa subjektia, kartesiolaista teatteria, jossa kokemukset ”esitetään” kokijalle (Dennett 1991, 39), ei ole myöskään yleisöä. Mutta illusiohan esitetään *jollekulle?*²⁵ Erityisesti Dennett on käsitellyt tätä argumenttia, ja hänen mielestään esityksen ”loppukäyttäjän” tuoli on syytä jättää tyhjäksi (Dennett 2016, 68). Illusionisti voi hyväksyä, että jokin sisäinen järjestelmä luo representaatioita, mutta sillä ei tarvitse olla katsojaa, varsinkaan tietoista. Kun representaatio on tehty, reaktio siihen voidaan delegoida erilaisille ei-tietoisille alijärjestelmille (Frankish 2016a, 34). Dennettin mukaan riittää, että järjestelmissä olevat sisäiset representaatiot ovat jotakin kognitiivisten kykyjen käytettävissä olevaa (representaatio käsi-silmä-koordinaation ohjaamiseksi, representaatio objektien tunnistamiseen jne.), eivätkä jotakin *meille*. Dennett pitää todennäköisenä, että representaation sisältöä ei voi kovin selvästi erottaa muista ominaisuuksista kuten siihen liittyvistä reaktioista ja assosiaatioista. (Dennett 2016, 68–69.)

Chalmersin mielestä vahva illusionismi ei ole epäkoherentti mutta empiirisesti väärin: kivun tuntemus *on aivan ilmeisesti olemassa*. Chalmers on jopa inspiroitunut G. E. Mooren skeptisismien vastaisen argumentin tapaiseen ”todistukseen”: Chalmers läimäisee itseään toiseen käteen ja toteaa tunteneensa kipua, joten kivun tuntemus on olemassa, m.o.t. (Chalmers 2018a, 53; 2018b, 53:29.)²⁶ Chalmersin argumentti kaikessa ytimekkydessään on seuraava:

- (1) Ihmiset tuntevat joskus kipua.
 - (2) Jos vahva illusionismi on totta, kukaan ei tunne kipua.
-
- (3) Vahva illusionismi on väärässä.

Vahvan illusionistin täytyy Chalmersin mukaan kieltää premissi (1) ja väittää että kivun

²⁵ Koska tämä intuitio on niin laajasti jaettu, Crick ja Koch (2003, 120) olettavat että se heijastelee jollakin tavalla aivojen yleistä rakennetta.

²⁶ Vrt. “If someone asks what conscious experience is, you say ‘Look, you know what is from your own case’ (If you want, you can add ‘Here’s an example’, and give them a sharp kick).” (Strawson 2017, 12.)

kokemus on vain introspektiivinen illuusio; jos vahva illusionisti myöntää, että ihmiset voivat *kokea* kipua, hän törmää taas tietoisuuden vaikeaan ongelmaan.²⁷ (Chalmers 2018a, 54.) Kuten seuraavassa luvussa käy selväksi, Dennett on valmis kieltämään premissin (1).

7 Dennett

Daniel Clement Dennett III on illusionisti, A-tyyppin materialisti, tunnettu tietoisuuden filosofian popularisoija ja polemisoiija ja Chalmersin pitkäaikainen antagonisti. Jyrkempää vastakkainasettelua kuin Chalmersin ja Dennettin välillä tuskin voi tietoisuuden filosofiasta löytää. Ennen kuin käsittelen Dennettin Chalmersiin ja muihin realisteihin kohdistamaa kritiikkiä, esittelen yksityiskohtaisemmin Dennettin aseman tietoisuuden filosofian kentällä.

Aluksi kuvaan Dennettin lähestymistavan, jota hän kutsuu *heterofenomenologiaksi*: käyn läpi tutkimuksen lähtökohdan ja metodin niin kuin Dennett ne artikuloi, ja selvitän hieman, minkälaisia tuloksia metodilla Dennettin mukaan pitäisi olla saatavissa. Sen jälkeen käyn lyhyesti läpi Dennettin tietoisuusteorian, huomio erityisesti siinä miten Dennett soveltaa viime vuosien aikana mielenkiintoa herättänyttä ennakoivan prosessoinnin teoriaa.

7.1 Heterofenomenologia

7.1.1 Tutkimusongelma

Tiede normaalisti tutkii ulkopuolelta, 3. persoonan näkökulmasta tutkimuskohdettaan. Tämä sallii havaittujen ilmiöiden käsittelyn näkökulmasta, joka sallii kontrolloidut toistokokeet, vertaisarvioinnin ja yhteisen kielen, jolla ilmiöistä puhutaan. Realistien argumenttien jälkeen vaikuttaa kuitenkin, että tietoisuuden kuvaamisessa tarvitaan jokin toinen kieli, yksityisen kokemuksen, 1. persoonan näkökulman kieli, ja kenties jopa aivan uudenlainen tiede. Dennett ei tätä hyväksy (ks. erit. Dennett 2001b), vaan uskoo, että perinteinen objektiivinen tiede on täysin ulotettavissa kattamaan myös tietoisuutta koskevat ilmiöt niin, että mitään ei jää selittämättä. Tätä tietoisuuden tiedettä Dennett kutsuu heterofenomenologiaksi (”hetero-”, eli fenomenologiaa *toisesta*), joka ”ottaa ensimmäisen persoonan näkökulman niin vakavasti kuin sen voi ottaa” (Dennett 2013,

²⁷ Dennett tuskin hyväksyisi tätä päättelyä. Hänen mukaansa ”kipu” on moninaisesti intuitioihin perustuva niin epämääräinen käsite, ettei sitä voi käyttää tällaiseen argumentointiin (Dennett 1978, 445).

341).

Descartes, samoin kuin Locke, Berkeley ja Hume kirjoittivat introspektion tuomia oivalluksiaan siinä käsityksessä, että lukijat voivat toistaa sanotut kokeet ja vahvistaa tutkimusten tulokset. Dennettin mukaan fenomenologia on perusteettomasti tukeutunut *1. persoonan monikkoon* (first-person plural), ajatukseen jonka mukaan voimme puhua yhdessä ensimmäisen persoonan kokemuksistamme²⁸ (Dennett 1991, 66–7).

Fenomenologia tarkoitti alun perin yksinkertaisesti ilmiöiden luettelointia, vasta 1900-luvulla se sai erityisesti tietoisuuteen liittyvän merkityksen Husserlin fenomenologiassa. Dennettin (Dan Zahavin mukaan kovin pintapuolisen²⁹) esityksen mukaan hankkeen menestys on jäänyt hyvin rajalliseksi ja fenomenologisia väitteitä koskevat kiistat ja ristiriidat ovat olleet ”nolostuttavan” yleisiä³⁰ (Dennett 1991, 67): ehkä olemme olettaneet liikaa toistemme samanlaisuudesta – tai ehkä introspektio ei ole niin luotettavaa kuin olemme olettaneet. Introspektio ei nimestään huolimatta ole vain ”sisään katsomista ja näkemistä”, siihen liittyy aina teoretisointia, Dennett sanoo. Me olemme erehtyväisiä, kuten tarkan näkökentän illuusio osoittaa, tai se että on olemassa ääni jonka sävelkorkeus tuntuu kohoavan loputtomiin ilman että se tosiasiallisesti muuttuu lainkaan. (Dennett 1991, 67–69; ks. myös Dennett 2003.)

Dennettin mukaan aivotutkijat ja kognitiotieteilijät ovat taipuvaisia siirtämään syrjään kysymykset tietoisuudesta ja keskittyvät tutkimaan ”ääreis-” tai ”alisteisia järjestelmiä”, joiden oletetaan jotenkin palvelevan jotakin hämärästi kuviteltua ”keskusta”, jossa ”tietoinen ajattelu tapahtuu”. Näin he vähättelevät sitä ymmärtämisen laajuutta, josta nämä ”ääreisjärjestelmät” ovat vastuussa, ja olettavat huomaamattaan kartesiolaisen teatterin – ja ”keskeisen merkityksenantajana” – olemassaolon. (Dennett 1984; Dennett 1991, 38–9, 228–9.) Kuin jossakin aivoissa olisi homunculus, joka seuraa monikanavaesityksenä aistien viestejä, lähettää käskyjä ruumiinjäsenille ja puhe-elimille, joiden tuottamille äänille vain tällä johtajalla on oikeus ja kyky antaa *merkitys*.

²⁸ Sen sijaan Max Velmans pitää tätä tarpeellisenä työhypoteesina: ”while *experienced* entities and events (phenomena) remain private to each observer, if their perceptual, cognitive, and other observing apparatus is similar, we assume that their experiences (of a given stimulus) are similar.” (Velmans 2007, 716.)

²⁹ Zahavi korostaa, että fenomenologia *ei ole* introspektiota: fenomenologia pyrkii etsimään intersubjektii-visesti tavoitettavia rakenteita ja sen analyysit ovat avoinna kaikkien fenomenologisesti asennoituneiden subjektien kritiikille (Zahavi 2007, 28 ja 31).

³⁰ Zahavin mukaan Dennettin vaatimus on liian kova: näillä ehdoilla useimmat teoriat, mukaan lukien Dennettin oma heterofenomenologia epäonnistuvat (Zahavi 2007, 37).

Dennettin mukaan tietoisuus on käyttäjäilluusio, ja me aliarvioimme suuresti niiden prosessien monimutkaisuuden, jotka saavat sen aikaan. Taikurit ovat, kuten Dennett sanoo, ”häpeämättömän viitseliäitä” rakentaessaan illuusioita, jotka nähdessään yleisö ei mitenkään voisi uskoa, miten pitkälle taikurit ovat valmiita menemään huijatakseen katsojia. Samoin luonto on häpeämätön – eikä sillä ole budjettirajoituksia mutta sen sijaan kaikki maailman aika käytettävissään (Dennett 2017, 31.)

7.1.2 Metodi

Esitän seuraavaksi heterofenomenologian metodin, niin kuin Dennett sen on kuvannut. Dennettin esittelee aluksi noudattamansa perussäännöt (Dennett 1991, 40):

1. Ei oleteta mitään nykyfysiikan ulkopuolista; ei vallankumousta materialismissa ennen kuin on aivan pakko.
2. Ei vain väitetä, että jotakin tietoisuuden ominaisuutta ei ole. Jos Dennett väittää jotakin tietoisuuden piirteen illusoriseksi, todistuksen taakka on hänellä.
3. ”Ei niuhoteta” empiiristä yksityiskohdista. Tieteen tarjoamaa nykytietämystä voi käyttää (empiiristesti kumottavissa olevien) teorialuonnostelmien muodostamiseen, vaikka yksittäisten tutkimustulosten merkitys on joskus kiistanalainen.

Jotta voidaan muotoilla teoria mielen sisällöistä, tarvitaan aluksi *neutraali tapa kuvailla havaintoja*. Pidättäydytään olettamuksesta tietoisuuden olemassaolosta esimerkiksi muilla ihmisillä (siis zombejakin voisi olla), mutta keskitytään ilmeisesti normaaleihin, aikuisiin ihmisiin, sillä jos jossakin on tietoisuutta, näillä sitä on. Kun näin on saatu ihmistietoisuuden teoria, voidaan jatkaa selvittämällä muiden olentojen mahdollisen tietoisuuden luonnetta. (Dennett 1991, 71–73.)

Tutkimuskohteena ovat tietoiset, hereillä olevat ihmiset, jotka voivat kertoa siitä ”millaista on olla” ja muista tuntemuksistaan sanallisesti, toimia yhteistyössä tutkijoiden kanssa ja suorittaa heidän määräämiään kokeita. Heterofenomenologia käyttää hyväkseen kykyämme *suorittaa* ja *tulkita* puheakteja. Ensimmäisenä havaintoaineistona ovat kuitenkin äännähdysten (tai esim. napinpainallusten) sarjat, joista tutkija tekee *litteroinnin*. Tässä askeleessa tehdään datan radikaali rekonstruktio (esimerkiksi) akustisesta muodosta sanajonoiksi. Tässä tapahtuu *tulkintaa*, joka perustuu tietoon siitä mikä kieli on kyseessä, että kieli on nimenomaan *ihmiskieli* jne. (Dennett 1991, 73–76.) Tuloksena saadaan luettelointi siitä, ”minkä subjekti uskoo olevan totta hänen tietoisesta kokemuksestaan”. Tämä luettelointi muodostaa tutkittavan subjektin *heterofenomenologisen*

maailman, ts. maailman niin kuin subjekti sen kokee. Tämä aineisto, yhdistettynä havaintoaineistoon subjektin aivotapahtumista ja ympäristöstä, erityisesti koejärjestelyistä, muodostaa ”kaikki mahdolliset objektiiviset ja subjektiiviset tietoisuuden ilmiöt”. (Dennett 2013, 342)

Tämän aineiston tulkitseminen vaatii *intentionaalisen asenteen* ottamista, ts. se vaatii työhypoteesin, jonka mukaan subjekti on *agentti*, jonka toimintaa ohjaavat rationaalisesti uskomukset ja halut ja muut mielentilat, jotka ilmentävät intentionaalisuutta ja jotka nekin ovat rationaalisia suhteessa subjektin kokemaan henkilöhistoriaan ja tarpeisiin (Dennett 1991, 76; 2013, 342–3). Litteroidun tekstin ajatellaan olevan *yhden, yhtenäisen subjektin* vilpittöntyä, luotettavaa ilmaisua tämän subjektin uskomuksista ja mielipiteistä (joskin erityisesti joidenkin patologioiden tapauksessa tämä etukäteisoletus voi olla ongelmallinen) (Dennett 1991, 77–78).

Tulkinnan sääntöjä voidaan eksplikoida, periaatteista, joiden mukaan intersubjektiivinen konsensus tuloksista saavutetaan voidaan neuvotella, subjektin rationaalisuuden luonnetta voidaan arvioida, ja kaikessa tässä noudatetaan normaaleja objektiivisen, empiirisen tieteen periaatteita (Dennett 2013, 343).

Tämä metodologia sitoutuu siis, paitsi fysikaalisesti havaittuihin ilmiöihin myös subjektin *uskomuksien* ja *halujen* olemassaoloon. Tutkijan on kuitenkin säilytettävä neutraali kanta, ja käsiteltävä näitä *teoreetikon fiktioina* tai *abstraktioina*, samantapaisina kuin painopisteet tai päiväntasaaja (Dennett 2013, 344).

Ne, jotka pitävät kokemusta episteemisesti ensisijaisena, väittävät, että kokemus itse on explanandum, eivät sitä koskevat raportit (esim. Levine 1994, 115). Dennettin mielestä tämä ei ole sopiva tieteellisen tutkimuksen lähtökohta. Hän esittää seuraavan luettelon aineistosta, järjestettynä vaaditun tulkinnan asteen mukaan (Dennett 2013, 344):

- a) ”tietoiset kokemukset itse”,
- b) näitä kokemuksia koskevat uskomukset,
- c) näitä uskomuksia koskevat sanallistetut arvostelmat,
- d) erilaiset äännähdykset ja toiminnat, joiden tulkitaan ilmaisevan em. arvostelmia.

Dennettin mukaan ulkopuolisen tutkijan näkökulmasta ensisijaista on tietenkin kohdan d) aineisto ja lisäksi mahdolliset aivosähkökäyrät, fMRI-skannausten tulokset jne. Tulkinnan tuloksena näistä saadaan aineistot c) ja edelleen b), mutta a):ta ei pidä yrittää sovittaa teoriaan. Voisihan olla, että 1) subjektilla on tietoisia kokemuksia, joita hän ei usko

hänellä olevan ja jolloin hänellä itselläänkään ei ole mitään pääsyä niihin, Dennett sanoo ja jatkaa: toisaalta voisi olla, että 2) subjektilla on uskomuksia tietoisista kokemuksista joita hänellä ei todellisuudessa ole, jolloin selitettävänä ovat subjektin uskomukset, eivät (olemattomat) kokemukset. (Dennett 2013, 344–5.) Jos subjekti yhä uskoo, että hänellä on jokin sanoinkuvaamaton kokemus, heterofenomenologi kuuntelee hänen mielipiteensä ja lisää tämän mielipiteen havaintoaineistoon: ”subjekti väittää, että hänellä on sanoinkuvaamaton kokemus” (Dennett 2013, 345).

Vaikuttaa hieman kummalliselta, miten kohdan 1) kaltaisia kokemuksia edes voisi olla: jos ne ovat tietoisia, subjektilla on varmaankin niitä kokeva uskomus. Samoin 2) vaikuttaa oudolta: voisiko minulla olla uskomus, että minuun sattuu, mutta todellisuudessa minulla ei ole kivun kokemusta? Mutta Dennettin metodologinen näkökulma on ymmärrettävä: jos a) sisältää havaintoaineistoa jota ei ole b):ssä, se ei ole tutkijan saavutettavissa. Ja toisessa tapauksessa mitään a):ta ei tietenkään olekaan.

7.1.3 Tulkinta

Se että jollekin käyttäytymiselle on koherentti tulkinta, ei takaa sitä, että tulkinta on myös oikea. Dennett toteaa, ettemme tosiaan voi olla varmoja siitä, että havaitsemamme puheaktit tosiaan ilmaisevat todellisia uskomuksia todellisista kokemuksista. Mutta yhdenkin intersubjektiiivisesti johdonmukaisen tulkinnan löytyminen on silti seikka, johon on syytä kiinnittää huomiota (Dennett 1991, 78).

Dennett vertaa heterofenomenologin tehtävää lukijan tehtävään tämän koettaessa tulkita fiktiivistä teosta. Riippumatta siitä tiedämmekö tarinan todeksi, voimme puhua siitä mikä on *tarinan sisäisesti totta*: kyllä, Sherlock Holmes todellakin asui Baker Streetillä. Tarinassa on myös totta paljon enemmän kuin siinä eksplisiittisesti kerrotaan: voimme sanoa tietävämmme, että Sherlock Holmesin Lontoossa ei lennellyt suihkukoneita, mutta että siellä oli pianonvirittäjiä, vaikka kumpaakaan ei eksplisiittisesti mainittaisi Doylen kirjoissa. Fiktiivisyydestä huolimatta romaanin maailmaa tutkimalla voi oppia paljon sen tekstistä, viestistä, kirjoittajasta ja jopa todellisesta maailmasta. Lisäksi, romaanin esittämästä maailmasta voi tietää paljon lukematta koko teosta (voin esimerkiksi tietää paljon Holmesin aikakauden Englannista, vaikka en olisikaan lukenut Doylen kirjoja). (Dennett 1991, 78–80.)

Kun tätä taktiikkaa sovelletaan tietoisuuden tutkimiseen, subjektin tuottama teksti generoi *teoreetikon fiktion* (Dennett 1991, 81): tutkija antaa tämän tekstin *konstituoida*

(fiktiivisen) maailman, subjektin heterofenomenologisen maailman, olettamatta mitään tämän fiktiivisen tarinan ja todellisen maailman välisestä suhteesta. Subjektin heterofenomenologinen maailma on vakaa, intersubjektiiivisesti vahvistettavissa oleva teoreettinen ehdotus, jolla on sama metafyyminen asema kuin esimerkiksi Sherlock Holmesin Lontoolla. Kuten fiktiossakin, se mitä sen (tässä tapauksessa mahdollisesti näennäinen) kirjoittaja sanoo, nauttii ehdotonta auktoriteettia (emme kysy, mistä Doyle tiesi Holmesin nojatuolin värin emmekä epäile, että hän ehkä tiesi sen väärin). Luotamme hänen (tulkittuun) sanaansa. (Dennett 1991, 81.)

Entä mistä kertoja todella puhuu? Tähän, Dennett sanoo, voi taas soveltaa vertausta romaanin tulkintaan: voimme analysoida, mitä kirjoittajan sukulaista kertomuksen henkilö muistuttaa, mitkä kirjoittajan lapsuudenkokemukset ovat päässeet kirjaan muodonmuutoksen kokeneina jne. Näiden analyysien jälkeen voimme esittää tulkinnan siitä, mistä romaani kertoo, jopa vastoin sen kirjoittajan väittämiä ja hänen kiivaita protestejään. (Dennett 1991, 84).

Vastaväitteisiin, että heterofenomenologia sivuuttaa *todelliset* tietoisuuden ongelmat, Dennett vastaa: heterofenomenologilla on viimeinen sana. Hän on tämän teoreetikon fiktion kirjoittaja. Jos haastateltava subjekti vaatii erehtymättömyyttä tietoisuutta koskeville väitteilleen, hän vaatii liikaa. Hän on erehtymätön vain siinä, mitä hänessä *näyttää* tapahtuvan, ja ulkopuolinen tutkija voi auttaa häntä testaamaan näitä väitteitä ja ehkäpä muokkaamaan niitä havaintoaineiston karttuessa. (Dennett 1991, 96; 2017, 351.)

Toisin sanoen, jos löytäisimme ihmisten aivoista ”todellisia” tapahtumia, joilla on tarpeeksi samoja määrittäviä ominaisuuksia kuin heidän heterofenomenologisella maailmallaan, voisimme perustellusti väittää, että olemme selvittäneet, mistä he *todella* puhuvat – ja jos huomaisimme, että nämä tapahtumat vastaavat tuskin ollenkaan subjektien heterofenomenologisia olioita, voisimme perustellusti väittää, että ihmiset ovat erehtyneet uskomuksissaan, vilpittömyydestään huolimatta (Dennett 1991, 85).

Ensi silmäyksellä Dennett näyttää ratkovan metodillaan tietoisuuden metaongelmaa: hän ei esimerkiksi yritä selittää, mikä synnyttää kvaliat, vaan mistä syntyvät niihin liittyvät uskomuksemme – paitsi että hän on tietenkin sitoutunut käsitykseen, että kvalioita ei ole (Dennett 1988; 1991, 372). Vaikka heterofenomenologi Dennettin mukaan sitoutuu ainoastaan uskomusten olemassaoloon ja pysyttelee neutraalina kysymyksessä fenomenaalisten kokemusten olemassaolosta, lopulta Dennett näyttää silti oletavan

ainoaksi todellisuudeksi näiden uskomusten virran (ks. esim. Zahavi 2007, 36). Tämä on Dennettin näkemyksen järkähtämätön perusta, joka asettaa hänet niin kauaksi Chalmersista.

7.2 Dennettin tietoisuusteoria

Pandemonium

Dennettin mallin mukaan erikoistuneiden, matalan taso prosessien, ”demonien” kuulumuus toisilleen mahdollistaa tietoisien tilan. Hieman samantapaiselle ajatukselle perustui myös Dennettin *monivedosmalli* (Dennett 1991, 101), jonka mukaan aivoissa virtaa rinnakkain useita narratiivisia ”vedoksia”. Monivedosmalli ei kuitenkaan tarjonnut ”tarpeeksi tehokasta vastalääkettä kartesiolaisen teatterin vetovoimalle”³¹, joten Dennett kehitti mielestään hyödyllisemmän metaforan *kuuluisuus aivoissa* (tai: *aivojulkis*; ”fame in the brain”, ”cerebral celebrity”) (Dennett 2001a, 224).

Pandemoniumissa ei-tietoiset sisällöt kamppailevat kuuluisuudesta keskenään ja tietoisuus käsittää juuri tämän kuuluisuuden piirin. ”Kuuluisuus” johtaa tietysti ajattelemaan, että täytyy olla katsojia, joiden havainnoinnin kohteena kuuluisuutta saavuttava on. Mutta ei ole olemassa kirjaimellista huomion valokeilaa, joten on selitettävä tämä funktionaalinen kyky huomion ”sieppaamiseen”, olettamatta mitään lähdettä joka ”suo huomiota” (Dennett 2001a, 225).

Ennakoiva prosessointi

Dennett vaikuttaa siirtyneen kannattamaan jonkinlaista *ennakoivaa prosessointia* (Dennett 2015a, 2015b), jota Wanja Wiese ja Thomas Metzinger kutsuvat ”kontrolloiduksi reaaliaikaiseksi hallusinaatioksi” ja jonka ajatellaan soveltavan bayesilaista päättelyä hyväkseen (Wiese & Metzinger 2017; ks. myös Seth 2016).

Perinteinen käsitys aivoista on ollut, että ne ovat passiiviset ja ulkoisen syötteen ohjaamat. Kuitenkin esimerkiksi havaittu taukoamaton hermoston aktiviteetti on johtanut näkemykseen, että aivot luultavammin itse asiassa *ennakoivat* tulevaa syötettä (A. Clark 2013, 2015). Väitteelle löytyy tukea myös keinotekoisien hermoverkkojen tutkimuksesta (esim. Rao & Ballard 1999). Tämä on vähemmän kuormittavaa aivoille, ja suurin osa informaation käsittelyn taakkaa johtuu *ennustusvirheistä*, jotka informoivat meitä siitä mikä on

³¹ Susan Schneiderin mukaan monivedosmalli lisäksi tukeutuu Bernard Baarsin globaalien työtilan malliin, joka taas perustuu juuri sellaiseen teatterimalliin, jonka Dennett haluaa torjua (Schneider 2007).

erilaista ja huomionarvoista (A. Clark 2015, 264). David Baßler huomauttaa, että tämä teoria olettaa selkeän erottelun mielen ja ulkomaailman välille (Baßler 2015, 415–6).

Ennakoiva prosessointi luo komputatiivisesti mahdollisimman taloudellisen representaation maailmasta, niin että aivojen tehtävä ohjata käyttäytymistä olisi mahdollisimman tehokasta, ja samaan tapaan me toimimme kun suuntaamme huomion itseemme ja omiin prosesseihimme (Wiese & Metzinger 2017).

Organismin toiminta taas ei ole reagoimista syötteeseen, vaan kätevä ja tehokas tapa valita seuraava ”syöte”. Järjestelmä ennakoi jatkuvasti tulevia tilojaan ja liikkuu itse tehdäkseen jotkin niistä tosiksi. Aistinelinten viestit tuottavat vain ennustuksia korjaavaa palautetta. (A. Clark 2013, 186; 2015, 264.) Motoriikalle ei tarvita erillisiä kustannuksia laskevia funktioita, ne absorboituvat ennustavaan järjestelmään (A. Clark 2015, 272). Motorinen ohjaus on *subjunktiivista*; se ennustaa, ei-aktuaalisia aistihavaintoja jotka *seuraisivat* jos toimisimme tietyllä tavalla. Ennustusvirheiden vähentäminen tapahtuu sitten tekemällä näistä tiloista aktuaalisia. (A. Clark 2015, 268–9.)

Havainnot esittävät meille ulkoisten objektien ominaisuuksia, vaikka todellisuudessa voimme havaita vain ulkoisten objektien vaikutuksia (Wiese & Metzinger 2017). Ulkoisen ilmiön havaittu vaikutus ei riitä määrittelemään yksikäsitteisesti sen syytä, joten joudumme tekemään hyviä arvauksia aiemman informaation perusteella. Jos meillä on informaatiota piilevistä syistä, voimme tehdä *ennustuksia* niiden vaikutuksista. Ennustusta verrataan toteutuvaan aistisignaaliin, tästä lasketaan *ennustusvirhe*, jonka mukaan lasketaan seuraava ennustus. (Wiese & Metzinger 2017.) Ja kun odotuksemme täyttyvät, ristiriidan puute toimii vahvistuksena ennustukselle.

Bayesilainen päättely on menetelmä, jolla todennäköisyyslaskennan periaatteiden mukaisesti yhdistetään olemassa olevaa informaatiota, joka on epävarmaa siinä mielessä että se voidaan esittää todennäköisyysjakaumina, uuteen havaintoaineistoon. Tällä tavalla saadaan usein tarkempi jakauma kuin ennen yhdistämistä. Bayesilainen päättely antaa rationaalisen ratkaisun ongelmaan, miten päivittää oletuksia havaintojen piilevistä syistä. (Wiese & Metzinger 2017.)

Phi-ilmio

Bayesilaisen päättelyn toimimisesta Dennett ottaa esimerkiksi phi-ilmion, niin kuin Paul Kolers ja Michael von Grünau (1976) ovat sitä testanneet: koehenkilölle näytetään peräkkäin kaksi hieman eri kohdassa sijaitsevaa eriväristä täplää, jolloin koehenkilö näkee

täplän liikkuvan kohti toista täplää ja vaihtavan väriään puolivälissä, *ennen kuin toinen täplä on näytetty* (Dennett 1991, 114). Miten se on mahdollista? Siten, Dennett selittää, että aivoilla on bayesilaisia odotuksia jatkuvasta liikkeestä, jotka luovat odotuksia väritäplien välisestä sisällöstä, ja jos tämän odotuksen kumoavaa informaatiota ei saavu, illusorinen havainto hyväksytään todelliseksi (Dennett 2015b, 406).

Dennettin toisessa esimerkissä, lauseessa ”Tom saapui Billin jälkeen” meillä ei ole ongelmaa ymmärtää järjestystä, vaikka nimet esiintyvät lauseessa ajallisesti käännettyssä järjestyksessä. Tässä kuten phi-kokeessakin koetun subjektiivisen ajan (ja tapahtumien järjestyksen) ei tarvitse olla yhtenevä tapahtumien objektiivisen ajan kanssa – eikä subjektiivinen aika ei ole mikään todella olemassa oleva osa tapahtuman kausaalista rakennetta. (Dennett 2015b, 407.) Tämän ajattelemisessa (kuten monien muidenkin ominaisuuksien kuten ”seksikkyyden” tai vitsin ”hauskuuden”) tarvitaan ”Humen outo inversio”, Dennett sanoo. Dennettin tarkoittama inversio koskee kausaliteettia, ja hyvin tyypistetysti ilmaistuna se tarkoittaa oivallusta, että kausaliteetin vaikutelma tulee *sisältä*, ei ulkoa. Näemme A:n, meidät on ”johdotettu” odottamaan B:tä – ja kun B tulee, uskomme havaintomme syyksi jonkin ulkoisen syyn, jonka uskomme havaitsemme jotenkin ”suoraan”. (Dennett 2017, 354–5.)

Kun Dennett näin on sijoittanut monien ominaisuuksien lähteen subjektiin, hän seuraa-vaksi vähän yllättäen kuitenkin muistuttaa, että havaittu väliväri phi-kokeessa, ”kvaliat”, ”hauskuus” ja muut tällaiset ominaisuudet eivät silti ole mitään ”intrinsisiä” tietoisien kokemuksen ominaisuuksia. Bayesilaiset odotukset luovat erilaisia odotuksia: odotamme kiinteillä, läpinäkymättömällä objekteilla olevan myös takapuolen, ovien aukenevan, portaiden mahdollistavan kiipeämisen ja niin edelleen. Mutta, Dennett sanoo, niiden olioiden joukossa, joilla on meille merkitystä, olemme myös *me itse*. Meillä pitää olla – ja meillä on – hyviä bayesilaisia odotuksia siitä mitä teemme, ajattelemme tai *odotamme* seuraavaksi. (Dennett 2017, 356–8.)

8 Dennettin kritiikki

Dennettin provosoiva ajatuskokeita, vertauksia ja anekdootteja pursuava tyyli tekee argumenttien seulomisen hänen teksteistään joskus hankalaksi. David Papineau (2017) toteaa, että vaikka Dennettin tekstejä on aina viihdyttävää lukea, hän jättää usein keskeiset käsitteensä sumuisesti määritellyiksi. Howard Robinson (1993) on kuvannut Dennettin metodia Jeriko-metodiksi: Dennett uskoo, että jos hän marssii tarpeeksi monta kertaa

filosofisen ongelman ympäri toivottaen tieteellisiä havaintoja, ongelma hälvenee silmiemme edessä – mutta varsinaiset filosofiset argumentit ovat harvassa.

Tulkintaa helpottaa kuitenkin, että hänen tietoisuuden filosofiansa perustavat vakaumukset ovat pysyneet lähes muuttumattomina ainakin 90-luvulta asti – jopa niin, että vastoin Dennettin väitteitä tieteelliset tutkimustulokset eivät näytä juuri vaikuttavan niihin³² (Papineau 2017). Merkittävimmit muutokset lienevät hänen siirtymisensä kannattamaan monivedosteorian sijaan pandemonium-mallia, ja aluksi kiinnostavana pitämänsä Rosenthalin korkeamman kertaluvun tietoisuusteorian hylkääminen.

Useimmat filosofiset teorit ovat vain *määritelmien puolustusta*, Dennett sanoo, ilman yritystä tehdä ennustuksia, ja näin ne, koska arkiuskomukset (folk intuitions) ohjaavat niitä, tuntuvat olevan aina askeleen jäljessä havaintoja. Illusionismia ei pidä nähdä yrityksenä kieltää ilmeinen vaan parhaana ehdokkaana kokemuksellisen tietoisuuden selittäjäksi, *oletusnäkemysenä* jota tulee pitää totena, kunnes toisin osoitetaan. (Dennett 2016, 67–68.)

Samalla hän tosin itse tuntuu hyväksyvän sellaisenaan oikeana tieteellisen 3. persoonan näkökulman, käsittelemättä sen pätevyiden ehtoja kuten esimerkiksi klassinen fenomenologia tekee (Zahavi 2007, 38–39). Tai sitä, että voidaan perustellusti väittää kaiken havainnoinnin olevan pohjimmiltaan yksityistä, ja että kolmannen persoonan tieto vaatii tulkitsevien subjektien olemassaolon ja on viime kädessä intersubjektiivista (MacLennan 1996, 410; Velmans 2007, 716–8). Tämä on tietenkin erityisen ongelmallista tietoisuuden vaikean ongelman tapauksessa, jossa ratkaistavana on paitsi mikä on hyväksyttävä lähtökohta, myös hyväksyttävä tutkimusmetodi.

Näin realistien ja illusionistien välisen kiistan vaikeus voidaan paikallistaa juuri näihin kahteen keskeiseen alkuolettamukseen.

Näiden huomautusten jälkeen siirryn käsittelemään Dennettin Chalmersin intuitioita kohtaan esittämää kritiikkiä, jonka ryhmittelen luvussa 5 esitetyn jaottelun mukaisesti. Tarvittaessa täydennän muiden kirjoittajien huomioilla. Koska Dennettin kritiikki tukeutuu paljolti hänen omaan tietoisuusteoriaansa, sivuan esityksessäni myös sitä.

³² “Dennett does consider a limited number of psychological experiments, but neurons, he tells us, ‘are not my department’ (pers. comm.)” (Crick & Koch 2003, 124.)

8.1 Metafyysiset intuitiot

Vaikea ongelma

Dennett kirjoittaa: ”Olen varsin varma, että houkutteleva ongelma jota nimitetään Vaikeaksi ongelmaksi, on yksinkertaisesti virhe. Mutta en pysty todistamaan sitä. Eikä Chalmers pysty todistamaan, että *on* olemassa vaikea ongelma.”³³ (Dennett 2013, 312 ja 316.) Dennett on nähtävästi myöntänyt, että väitteessä, että fyysikaalisen lisäksi on jotakin joka vaatii selittämistä, ei ole mitään loogisesti ristiriitaista.³⁴ Dennett epäilee kuitenkin, että selvittäessämme ”helppoja ongelmia”, tutkiessamme jotakin tiettyä ratkaisematonta ongelmaa meiltä unohtuvat helposti toiset, jo ratkaistut ongelmat, ja niinpä meille jää jokin vaivaava tunne *niistä muista kysymyksistä*, jostakin jäänteestä, joka sitten ilmeisestikin kuuluu vaikean ongelman piiriin (Dennett 2013, 316–7).

Dennettin käsittää tietoisuuden paitsi aivorakenteen synnyttämäksi myös kokoelmaksi meemejä, ja tietoisuuden olemassaolo riippuu vahvasti tietoisuuteen liittyvistä käsitteistä (Dennett 1991, 24). Block kritisoi tätä todeten, että aivan ilmeisesti fenomenaalinen tietoisuus *itse* ei ole kulttuurinen konstruktio, toisin kuin sitä koskeva käsite; kulttuuri kuitenkin *vaikuttaa* fenomenaaliseen tietoisuuteen, mistä todisteeksi käy esimerkiksi viinin makuaistimuksen kehittyminen (Block 1995, 238).

Dennettin mielestä bayesilaisten odotusten vuoksi me luulemme introspektiossa havaitsevamme ”intrinsisiä” ”lisä”ominaisuuksia. Tapahtuu jonkinlainen Humelainen projektio ”subjektiiviseen avaruuteen” (Dennett 2015b, 406). Mutta ei ole mitään subjektiivista avaruutta missä tämä tapahtuu eikä näitä ominaisuuksia todella olemassa olevina. Dennett on myös koettanut korostaa vaikean ongelman illusorisuutta vertaamalla sitä ”rahaan”, ”rakkauteen” tai ”söpöyteen” (esim. Dennett 1996, 6), todeten että ne ovat samankaltaisia kuin tietoisuus: mahdollisesta ensivaikutelmasta huolimatta niille ei tarvitse olettaa mitään itsenäistä olemassaoloa, eikä ole mitään syytä postuloida esimerkiksi söpöyttä fundamentaaliseksi ominaisuudeksi fysiikkaan massan, varauksen ja aika-avaruuden lisäksi. Samoin Dennett on verrannut fenomenaalista tietoisuutta myös vitalismiin³⁵

³³ Aiemmin Dennett nimitti Chalmersin vaikeaa ongelmaa teoreetikon fiktioksi (Dennett 1996, 1998), joka ei ole todellinen ongelma, joka vaatisi vallankumouksellista uutta tiedettä (Dennett 2001a, 233).

³⁴ Patricia Churchlandin mukaan ”vaikea ongelma” suojaa itsensä ”tavaramerkin oikeuksilla”, jonka jälkeen kaikki sen selitysyrietykset ovat välttämättä vääriä (P. S. Churchland 1996; ks. myös Carruthers & Schier 2017).

³⁵ Vastaavia vertauksia ovat esimerkiksi Paul Churchlandin (1996) ajatuskoe filosofista joka miettii ”valon vaikeaa ongelmaa” ja Patricia Churchlandin (1996) vertaus keskiaikaiseen käsitykseen ”kuunylisen” maailman mekaniikan selittämättömydestä. Chalmersin vastaväite näihin on samankaltainen kuin söpöyden ja vitalismin tapauksessa (ks. esim. Chalmers 1997, 4–6).

(Dennett 1991, 281–2; 1996, 1).

Chalmers ei pidä näitä argumentteja pitävinä: esimerkiksi vitalistin ei kaikkien elämän funktioiden selittämisen jälkeen olisi edes *käsitteellisesti* mahdollista väittää, että on vielä jotain selittämistä – kun taas tietoisuuden tapauksessa oma kokemuksemme antaa meille *riippumattoman syyn* olettaa, että on vielä jotakin muutakin selitettävää kuin psykologiseen tietoisuuteen liittyvät funktiot (Chalmers 1996, 109; 2002a, 252).

Luonnollisesti Dennettin mielestä näiden funktioiden selittämisen jälkeen ei jää enää mitään selitettävää (Dennett 1996, 5).

Tietoisuus syntyy monimutkaisuudesta

Dennett väittää hyökkäyksen reduktiivista functionalismia vastaan perustuvan yksinkertaiseen päättelyvirheeseen: on selvää, että millään ohjelmapätkällä, joka on tarpeeksi lyhyt että voimme helposti ymmärtää sen, ei ole aitoa ymmärryskykyä, ja niinpä päättelemme, että millään määrällä tällaisia ohjelmapätkiä ei voi olla aitoa ymmärryskykyä – mutta miksi tämä olisi totta? (Dennett 1991, 438–9.) Monimutkaisuudella on merkitystä. Jos ei olisi, meillä olisi paljon yksinkertaisempi argumentti vahvaa tekoälyä vastaan: ”Katso, tämä laskin ei ymmärrä kiinaa, ja jokainen ajateltavissa oleva tietokone on vain suuri laskin, joten mikään tietokone ei voi ymmärtää kiinaa.” Kun otamme huomioon monimutkaisuuden, meidän täytyy *todella* ottaa huomioon se eikä vain teeskennellä, että otamme. Esimerkiksi Searlen (1980) kiinalaisen huoneen ajatuskoe, aivan kuten Jacksonin Mary-ajatuskoe, tuottaa vahvan vakuuttuneisuuden ”vain kun emme seuraa ohjeita”. (Dennett 1991, 440.)

Dennettin ja Doug Hofstadterin mielestä Searlen ajatuskoe – jota Dennett kutsui *intuitiopumpuksi*, joksikin joka on suunniteltu tuottamaan intuitiivisia mutta vääriä vastauksia (Dennett 1991, 282) – oli paitsi kiehtova myös virheellinen ja harhaanjohtava. Miettiessään miten Kiinalaisen huoneen argumentti toimii, Hofstadter esitti, että heidän tulisi ”kääntää kaikkia säätimiä” ajatuskokeesta nähdäkseen mikä saa sen toimimaan³⁶. (Dennett 2013, 320) Tässä tapauksessa oikea säädin oli Dennettin mukaan se, joka sääteli kiinalaisen huoneen mekanismien monimutkaisuutta. Oli ilmeistä, että järjestelmällä pitäisi olla kyky käsitellä lauseita, paljon arkitietoa maailmasta, ohjeita miten reagoida piloihin, vihjailuihin jne. Mutta tällöin ei olisikaan enää intuitiivisesti selvää, että Searlen

³⁶ Jaakko Hintikalla on vastaava esimerkki intuition muuttumisesta kun ilmeiseksi väitetyn yksittäistapauksen parametreja muutetaan: ”prima facie intuitions do not come fully equipped with instructions for their use.” (Hintikka 1999, 137–8.)

ajatuskoe toimisi ajatellulla tavalla (Dennett 2013, 323–9). Kiinalainen huone on antanut tukea intuitiolle, että prosesseista ja prosessien kokonaisuuksista, jotka voi kuvata funktionaalisesti, ei koskaan voi syntyä *tietoisuutta* – mutta jos ajattelemme lukemattomien mekanismien ”päälle rakentuneiden virtuaalikoneiden päällä olevien virtuaalikoneiden päällä olevia virtuaalikoneita” (Dennett 2013, 325), ei enää vaikutakaan yhtä ilmeiseltä, ettei tietoisuus voisi syntyä niistä. Dennett sanoo Searlen intuitiopumppua ”epäkuntoiseksi”, joksikin joka on rakennettu pysäyttämään intuition johtuminen tiettyyn (tässä tapauksessa monimutkaisuuden) suuntaan (Dennett 2013, 320).

Ymmärrys ei ole jumalallinen taito, josta kaikki suunnittelu virtaa, vaan emergentti ilmiö joka nousee ei-ymmärtävistä osajista (Dennett 2017, 75). Dennett kirjoittaa *taitavuudesta ilman ymmärrystä* (competence without comprehension) viitaten Darwinin ja Turingin oivalluksiin. Darwin käsitti, että moitteettomasti toimivan (biologisen) koneen tekemiseksi ”ei ole tarpeen tietää, miten se tehdään” (Dennett 2009; 2017, 53). Turing taas oivalsi, että universaalien tietokoneiden riittää toimia täysin mekaanisesti pienin, ei-älykkäin askelin (Dennett 2017, 55).

”Heureka-hetket” näyttävät todistavan, että ymmärrys on jonkinlainen *kokemus*, mikä on johtanut jotkut väittämään, että ymmärtäminen ei ole mahdollista ilman tietoisuutta (erit. Searle 1992). Mutta Dennett esittää, että eläinaivojen (mukaan lukien ihmisaivojen) perusarkkitehtuuri koostuu luultavasti bayesilaisista verkostoista, jotka ovat hyvin päteviä odotusten generoijia, niin että niiden ei tarvitse ymmärtää, mitä ne tekevät. Mallina tälle Dennett käyttää havaintoa ”syväoppimisesta” ja bayesilaisilla menetelmillä toimivista järjestelmistä, huomioiden miten hyvin ne toimivat ja miten hyviä vastauksia ne antavat vaikeisiin kysymyksiin ”vaikka ne eivät kykene kertomaan meille miten ne sen tekevät”, toisin kuin Haugelandin (1985) eksplikoima GOFAI. (Dennett 2017, 94–95, 174–5, 316.)

Fenomenaaliset ominaisuudet voisivat olla ”struktuuria alas asti”, jolloin ne sittenkin voisivat olla redusoitavissa fysikaalisiin ominaisuuksiin. Chalmers on pohtinut tätä mahdollisuutta, ja hänen ratkaisunsa tähän on löytää redusoitumattomuus siitä, miten fenomenaaliset ominaisuudet intrinsisinä ovat riippumattomia kaikista aina relationaalisesti ilmaistavissa olevista fysikaalisista käsitteistä (Chalmers 1997, 21). Tämä väistöliike ei kuitenkaan auta Chalmersia, jos käsitys fenomenaalisten ominaisuuksien intrinsisydestä osoittautuu kestävämmäksi, niin kuin se jäljempänä paljolti osoittautuu.

Aiemmin mainittu Bourgetin ja Chalmersin kyselytutkimus heikentää myös metafyy-sisten intuitioiden asemaa: jos filosofien intuitiot jakautuvat niin kuin ne jakautuvat, vaikean ongelman ilmeisyys explanandumina on kyseenalainen. Toinen samalla tavoin merkittävät tulos tulee Justin Sytsman ja Edouard Macheryn (2010, 2012) empiirisistä tutkimuksista, joiden mukaan tavalliset ihmiset eivät koe subjektiivisilla kokemuksilla olevan fenomenalisia ominaisuuksia, toisin kuin filosofit – minkä lisäksi filosofit eivät vaikuta olevan tietoisia tästä eroavaisuudesta! Filosofien intuitio ei myöskään vaikuta olevan ratkaisevasti luotettavampaa kuin filosofisesti kouliintumattomilla ihmisillä (Weinberg, Gonnerman, Buckner, & Alexander 2010).

8.2 Selittävät intuitiot

Chalmers sanoo, että ”tietoisuus näyttää olevan lisä fakta tietoisista järjestelmistä”. Dennett on samaa mieltä siitä, että siltä näyttää. Tämä on mitä Dennett on kutsunut *zombivainuksi* (zombic hunch). Dennett sanoo tuntevansa sen yhtä elävästi kuin kuka tahansa muukin – hän ei vain myönnä sille käypää arvoa, sen enempää kuin usein lähes vastustamattomalta tuntuvalle tunteelle, että Aurinko kiertää Maata: se totisesti näyttää kiertävän. (Dennett 2012, 88; 2013, 283–5).

Chalmersin organisatorisen invarianssin periaate on hyvin lähellä laadullisten ominaisuuksien identifioimista funktioihin. Periaate esittää lisäksi todella yleisen lainalaisuuden, jonka Chalmers postuloi jokseenkin perustelematta. Sitä tukemaan tarkoitettut hiipuvien ja tanssivien kvalioiden argumentit eivät nekään vaikuta välttämättä pitäviltä (Greenberg 1998).

Niinpä Dennettillä on hyvä syy ihmetellä, miksei Chalmers ole A-tyyppin materialisti, kun Chalmers luo Dennettin mielestä erinomaisen argumentin materialismin puolesta: Chalmers pohtii tietoisuuden lataamista muistipankkiin ja pitää outona ajatusta, että fenomenaalinen tietoisuus voisi himmentyä tai kadota jos ladattu ihminen olisi funktio-naalisesti identtinen (Chalmers 2010, 46; Dennett 2012, 89–90). Chalmers perustelee, että on loogisesti mahdollista että tietoisuus ei säilyisikään siirrossa. Dennett luettelee joukon loogisesti mahdollisia outoja asioita ja toteaa, ettei pelkän loogisen mahdollisuuden perusteella kannata hylätä lupaavaa tietoisuuden teoriaa. (Dennett 2012, 90; ks. myös 1991, 430.)

Sanottakoon kuitenkin, että Chalmers puhuu säilymisen yhteydessä *luonnollisella* välttämättömyydellä säilyvistä ominaisuuksista, ts. ei-reduktiivisen funktionalismin

näkemyksestä. Jos Chalmers luopuisi fysikaalisista tietoisuuden *emergoitumisesta* psykofyysisten lakien avulla, hänen argumenttinsa todellakin tukisi Dennettia.

Selityksellisen kuilun voi sulkea osoittamalla, että kvalioiden yksityisyys ja sanoinkuvaamattomuus selittyvät täysin funktionaalisesti, ja että niiden intrinsinen, kvalitatiivinen luonne ei ole lainkaan olemassa eikä näin tarvitse selittämistäkään (T. W. Clark 1995, 252). Tässä Dennett hyvin pitkälle onnistuuakin kvaliakritiikissään.³⁷

Chalmersin mukaan Dennett, kuten A-tyyppin materialistit yleensä, vain *olettaa* että pelkästään funktiot tarvitsee selittää, eikä argumentoi olettamuksensa puolesta juuri ollenkaan. (Chalmers 1997, 7.) Tietenkin Chalmersia itseään kohtaan voi esittää vastaavanlaisen kritiikin: hänen alkuoletuksenaan on, että kokemus ei ole identtinen fysikaalisen kanssa (esim. T. W. Clark 1995, 242).

Selityksellisen kuilun yhteydessä on sivuttava hiukan B-tyyppin materialisteja. Heidän argumenteistaan tavallinen on fenomenaalisten käsitteiden strategia: selityksellinen kuilu on olemassa koska fenomenaaliset käsitteet ovat erityyppisiä kuin fysikaaliset tai funktionaaliset käsitteet, vaikka fenomenaaliset ominaisuudet ovatkin identtisiä joidenkin fysikaalisten tai funktionaalisten ominaisuuksien kanssa (Chalmers 2006, 167–168). Esimerkiksi Papineau (2006) mielestä fenomenaalisen käsitteen yhteydessä täytyy esiintyä todellinen tai uudelleen kuviteltu instanssi viitatusta fenomenaalisesta tilasta, kun taas identiteetin fysikaalista edustavalla puolella ei, minkä vuoksi identiteetti vaikuttaa epätodelta.

Chalmersin mukaan B-tyyppin materialistilla on silti vaikeuksia selittää fenomenaalisten käsitteiden erityisluonne materialismin kanssa yhteensopivasti. Fenomenaaliset käsitteet ovat joko *paksuja*, jolloin niitä ei voi selittää fysikaalisesti – kuten ehdotuksessa, että fenomenaaliset käsitteet sisältävät suoran russelilaisen tuttuusrelaation, samankaltaisen kuin minkä Chalmers (2003) esittää. Tai ne ovat *ohuita*, jolloin ne voidaan kyllä selittää fysikaalisesti mutta ne eivät kykene selittämään episteemistä tilannettamme. (Chalmers 2006, 180–189.) Ja lisäksi tällaiset identiteetit ovat silloittavia periaatteita fysikaalisesta fenomenaaliseen, jotka B-materialisti joutuu olettamaan fysikaalisten tosiasioiden lisäksi (Chalmers 1997, 12; 2002a, 255; tämä tutkielma, luku 4.2).

³⁷ Ks. luvut 8.6 ja 8.7.

8.3 Tietointuities

Lepakoista

Dennettin mukaan, kun saamme heterofenomenologisen kertomuksen jonka hylkäämiselle ei ole mitään perusteita, meidän pitäisi hyväksyä se – alustavasti, riippuen myöhemmistä löydöistä – täsmälliseksi kuvaukseksi siitä, millaista on olla kyseinen olento. Niin me kohtelemme toisiammekin. (Dennett 1991, 443–4.)

Mutta nämä tutkimukset eivät kerro, *mistä* lepakot ovat tietoisia. ”Ehkä ne ovat lentäviä zombeja?” Dennett parodioi, ja jatkaa: ehkä lepakko vain automaattisesti rekisteröi doppler-siirtymiä sisäisellä tietokoneellaan (kuten poliisi tutkallaan) ja yhtä automaattisesti hypähtää lentoon? Paitsi jos siellä on jokin sisäinen tarkkailija, joka reagoi sisäiseen esitykseen paljolti kuten poliisi reagoi tutkansa nopeusnäyttöön. Mutta tämä on vain vanha kartesiolaisen teatterin harha: siellä *ei* ole representaatiota eikä sisäistä tarkkailijaa. (Dennett 1991, 444–5.)

Lepakon tietoisuudesta on vaikea saada tietoa, koska lepakot eivät puhu. Dennettin mielestä silti heterofenomenologia ilman tekstiä on, vaikkakin vaikeaa, kuitenkin mahdollista. Lepakon mielen rakenne on yhtä saavutettavissa kuin sen ruuansulatusjärjestelmäkin. (Dennett 1991, 446–7.) Oivallukset, joita voi saada ensimmäisen persoonan tarkkailusta on alistettava kontrolloituihin kokeisiin naiivien subjektien kanssa. Tällä tavalla varmistettu ensimmäisen persoonan kokemuksesta saatu tieto sopii hienosti yhteen heterofenomenologian kanssa. Mutta ”löydöt”, joita ei voi testata tällä tavalla, eivät hyväksyttävää havaintoaineistoa (Dennett 2001a, 230).

Muutosokeus

Dennett argumentoi, että ihmisiltä voi jäädä huomaamatta toistuvasti muutos kuvan väreissä, kunhan vain kuvien näyttämisen välillä on lyhyt peittävä ärsyke (masking stimulus). Värimuutokset tulevat rekisteröidyiksi jollakin tasolla subjektin näköjärjestelmässä, mutta ovatko ne esillä visuaalisessa fenomenologiassa? Realisteilla on tässä dilemma. Joko he sanovat, että fenomenaaliset ominaisuudet muuttuivat heidän huomaamatta sitä, jolloin he joutuvat myöntämään, ettei heillä ole erityispääsyä näihin ominaisuuksiin ja ne saattavat muuttua koko ajan heidän tietämättään. Tai sitten heidän täytyy sanoa, että fenomenaaliset ominaisuudet muuttuivat vasta kun he huomasivat muutoksen – eli nähtävästi fenomenaaliset ominaisuudet ovat vain arvostelmiemme luomia rakennelmia, juuri kuten illusionistit väittävät. Dennettin loppupäätelmä on, että

”fenomenaalinen ominaisuus” on vain sekava käsite omiaan tuottamaan hämmennystä. (Dennett 2005, kappale 4.) Sen sijaan, yhteensopivasti Dennettin näkemyksen kanssa, ennakoiva prosessointi näyttäisi soveltuvan muutossokeuden selittämiseen (Madary 2015, 290).

Arkipsykologisen mallin kritiikki

Arkipsykologian mukaan subjektin tietoinen elämä on yksityistä. Subjekti voi valita, mistä ”sisäisistä kokemuksistaan” hän päättää kertoa, jonka jälkeen hän muotoilee verbaalisen ilmauksensa vastaamaan mahdollisimman tarkasti kokemuksensa sisältöä. Mutta tämä sanojen valintaprosessi ei ole ymmärretty: me vain hyväksymme joitakin sanoja ja hylkäämme toisia, ja perustelemme *jälkikäteen* valintamme niin, että ne näyttävät ilmeisiltä (Dennett 1991, 304).

Realisti väittää: ”Voin tehdä virheen kertoessani tietoisista tiloistani, mutta silti on olemassa *totuus* näistä tiloista.” Dennettin mukaan väitteessä on monia välivaiheita: 1) kokemukseni, 2) uskomus että minulla on tuo kokemus, 3) tähän uskemukseen liittyvä ajatus, 4) intentio ilmaista se, ja 5) puheilmaisu. Missä kohtaa vain voi sattua virhe, jolloin ilmaisemme jotakin muuta kuin mitä koemme. Niinpä tulee kiusaus ”oikaista”, ja julistaa, että ajatukseni (tai uskumukseni) siitä miltä kokemukseni *näyttää* minulle, on juuri mitä minun kokemukseni *todella* on. (Dennett 1991, 317–8.) Tässä on Dennettin vastaväite intuitioon, jonka mukaan miltä kokemus näyttää ja mitä se todella on, ovat sama asia.

Myös muistoni voi olla vääristynyt: vaikka osaisin sanoa, miltä minusta tuntuu nyt, voin hyvin erehtyä siitä, miltä minusta tuntui aiemmin – vaikka viive olisi lyhytkin. Enkä voi tietää, onko se vääristynyt *orwellilaisesti* (sinne on lisätty jokin fakta jota kokemishetkellä siellä ei ollut) vai *stalinilaisesti* (kokemus on alun perin väritynyt virheellisesti jonkin aiemman muiston vuoksi). Eli vaikka minulla olisikin ”suora” ja ”välitön” pääsy arvostelmaani, tämä saattaa olla vain virhearvostelma siitä, miltä minusta vaikutti hetki sitten (Dennett 1991, 115–26, 318.) Subjektista näyttää, että hänellä on sekä uskumuksia kokemuksistaan että lisäksi kokemukset itse, mutta Dennettin näkemyksen mukaan nämä molemmat ovat osa sitä, miltä asiat heistä näyttävät, ja osa subjektin heterofenomenologista maailmaa joka on selitettävä (Dennett 1991, 391).

Chalmersin mielestä Dennett harppaa siitä, että hänen teoriansa selittää miksi tietoisuuden sisällöt näyttävät meille siltä kuin näyttävät, siihen, että hänen teoriansa selittää kaiken mikä tarvitsee selittää (Chalmers 1996, 190). Chalmersin mukaan Dennett käyttää

hyväkseen sanan ”näyttää” monimerkityksisyyttä. On olemassa ”näyttämisen” 1. fenomenaalinen merkitys, jossa ”joltakin näyttäminen” on sen *kokemista* tietyllä tavalla, ja 2. psykologinen merkitys, jossa ”joltakin näyttäminen” tarkoittaa, että olemme taipuvaisia *muodostamaan arvostelman*, että ne ovat niin. Kokemuksen teorian pitäisi selittää miltä asiat näyttävät 1. merkityksessä, mutta Dennettin teoria selittää sen 2. merkityksessä (Chalmers 1996, 190).

Mary

Dennettin mielestä Mary-ajatuskokeessa tehdään samantapainen virhe kuin Searlen kiinalaisen huoneen ajatuskokeessa: ei ajatella mitä ajatuskokeen asetelma todella merkitsee. Mutta oletetaanpa että Marylla tosiaan on, ei huomattavan paljon fysikaalista tietoa, vaan *kaikki* fysikaalinen tieto. Sitten hän pääsee ulos huoneesta ja hänelle näytetään banaania ja kysytään, mikä värinen se on. Maryn tietämättä tutkijat ovat tehneet käytännön pilan ja maalanneet sen siniseksi. Silti Mary vastaa erehtymättä: ”se on sininen”. Tutkijat ällistyvät: miten se on mahdollista? Mary selittää, että koska hänellä on täydellinen fysikaalinen tieto, hän totta kai myös tietää, miten sininen ja keltainen valo vaikuttavat hänen silmänsä verkkokalvoon ja tietää, millainen vaikutus niillä on hänen kognitiiviseen järjestelmäänsä. (Esim. Dennett 1991, 398–406.)

Tämä vasta-argumentti ei kuitenkaan toimi. Siitä, että Mary tietää kaiken fysikaalisesta, ei seuraa, että hän on *kaikkietävä*. Marylla voi hyvin olla kaikki yleinen fysikaalinen tieto, mutta hän ei silti voi tietää *vain katsomalla*, mikä fysikaalisen ärsykkeen banaani tuottaa hänelle – jollei hänellä jo ennakolta ole kykyä yhdistää kokemuksellinen väri hänen omaan fysikaaliseen tilaansa (esim. Robinson 1993). Lisäksi vaikka Mary kykenisi tähän, sama tieto olisi Marylla myös ennen huoneesta pääsyä, eikä se silti kertoisi Marylle, millaista on *kokea* punaista (Chalmers 1996, 145).

Robinsonin kritiikki ei vakuuta Dennettiä: hänen mukaansa Maryn värikokemuksessa ei ole mitään sanoinkuvaamatonta tai analysoimattomissa olevaa, se vain on *hyvin* informaatorikas (Dennett 2005, 111). Dennett esittelee RoboMaryn, joka on konekopio Marystä ja jolla alkuperäisen oletuksen mukaan on moitteeton värijärjestelmä mutta vain puuttuvat värikamerat ja lukitus joka estää häntä muokkaamasta täydellisellä tiedollaankaan tilarekistereitä joihin ”värikokemus” vaikuttaisi. RoboMary kiertää tämän ongelman rakentamalla varamuistiinsa täydellisen kopion itsestään ilman näitä rajoitteita, jonka jälkeen hän pystyy kopioimaan kloonatun RoboMaryn värikokemuksen *aiheuttaman tilan* rekisteriarvot itselleen ja tietää silloin täsmälleen, millaista on nähdä punaista.

(Dennett 2005, 122–129.)

RoboMarya kuitenkin voi tuskin enää sanoa funktionaalisesti samanlaiseksi kuin alkuperäistä Marya, joten tämänkin argumentin pätevyys on vähäinen.

Epifenomenaaliset kvaliat

Jacksonin mukaan näkökokemuksissa on kvalioita, jotka ovat epifenomenaalisia eli niillä ei ole lainkaan vaikutuksia fysikaaliseen maailmaan. Silloin niitä ei voi mitenkään havaita, eivätkä ne itse voi toimia todisteena olemassaolostaan – joten mitä syytä on olettaa niiden olemassaoloakaan (Dennett 1991, 400–2; Frankish 2016a, 24)? Tietenkin voisi väittää, että epifenomenaalisilla kvalioilla on vaikutus subjektin ei-fysikaalisiin uskomuksiin, muuta sitten *näillä* uskomuksilla ei voisi olla mitään vaikutusta fysikaalisessa maailmassa. (Dennett 1991, 403.)

Epifenomenaalisesta toisesta, ei-filosofisesta merkityksessä se on epäolennainen oheistuote, kuten jalan naputtaminen lukiessa, jolla toki voi olla vaikutuksia fysikaaliseen maailmaan (Dennett 1991, 401–2). Tietoisuus on siis epifenomenaalinen joko filosofisessa, ”naurettavassa” mielessä (Dennett 1991, 405) – Dennett ei perustele tätä sen enempää – tai ei-filosofisesta mielessä, jolloin vaikka zombit ja ihmiset ehkä voitaisiin erottaa jonkin fysikaalisen eron avulla, tietoisuus ei näy funktionaalisenä erona tarkkailijalle. Tämäkin on ”naurettavaa”, sillä mikään ei kiinnitä tietoisuutta mihinkään tiettyyn fysikaaliseen eroon, ja valinta on mielivaltainen. (Dennett 1991, 405.)

Dennett törmää tässä samaan vaikeuteen kuin vaikean ongelman yhteydessä: hän myöntää, ettei voi todistaa, että epifenomenaalista tietoisuutta ei ole. Joten paras mitä hän voi tehdä, on osoittaa, ettei siihen uskomiseen ole mitään hyvää perustetta (Dennett 1991, 406).

8.4 Modaali-intuitiot

8.4.1 Zombit ja invertoidut kvaliat

Zimbot

Keskustelussa Dennettin kanssa Chalmers on myöntänyt, että hänen intuiotensa koskien zombien mahdollisuutta on perusteeton mutta myös horjumaton, joten Dennettin antamaa neuvoa terapiasta tai ruokavalion vaihtamisesta (Dennett 2012, 91) voi tässä yhteydessä pitää jopa filosofisesti perusteltuna. Dennettin sanottava ei tietenkään lopu tähän.

Kun filosofit väittävät, että zombit ovat kuviteltavissa, he aliarvioivat taas tarvittavan

mielikuvituksen määrän (Dennett 1995, 322). Tätä valaistakseen Dennett esittelee ajatuskokeen *zimbosta*³⁸, joka on kuten zombi, mutta sillä on sisäisiä korkeamman tason (eietoisia) informaatiotiloja, jotka koskevat sen alemman tason informaatiotiloja. Zimbolla on kaikesta mitä se tekee korkeamman tason representaatioita, se voi raportoida tiloistaan, läpäistä Turingin testin, mutta kaikki sen toiminnot ovat tiedostamattomia. Zimbo uskoo (tiedostamattomasti) että se on erilaisissa mentaalisissa tiloissa, että se on tietoinen. Sillä on siis Chalmersin kuvailemia ja psykologiseen tietoisuuteen kuuluvia itsereflektiivisiä funktioita. (Dennett 1991, 310–1; 1995, 323).

Tällainen zimbo olisi Dennettin käsityksen mukaan tietoinen, koska se kykenee tekemään kaiken sen kognitiivisen työn joka kuuluu realistien olettamalle tietoisuudelle (Dennett 1995, 322–4). Ja ennen kaikkea, näillä käsitteillä me kaikki olemme zimboja. Samalla tavalla kuin me, zimbo olisi myös oman virtuaalikoneensa käyttäjä-illuusion vallassa (Dennett 1991, 311; Dennett 2013, 290), vaikka illuusiota ei tietenkään esitetä kenellekään; vertaus tietokoneen käyttöliittymään toimii vain tiettyyn rajaan asti.

Invertoidut kvaliat

Realistien käyttämään invertoitujen kvalioiden argumenttiin Dennett kehittää seuraavan ajatuskokeen. Oletetaan että on jokin laite, joka ottaa syötteenä esimerkiksi Jorma Uotisen aistimukset ja syöttää ne minun päähäni. Nyt pystyisin silmät kiinni raportoimaan kaiken mitä Uotinen näkee – mutta mitäpä jos nyt näenkin taivaan keltaisena, ruohikon punaisena jne.? Tämä vaikuttaisi vahvistavan empiirisesti, että meidän kvaliamme ovat erilaiset. Mutta kuvitellaan sitten, että teknikko irrottaa kaapelin pistokkeen, kääntää sitä 180 astetta ja laittaa takaisin, ja nyt näen taivaan sinisenä, ruohikon vihreänä jne. Oliko ensimmäinen kaapelin asento väärä, vai oliko se oikea ja Uotisen aistimukset ovat erilaiset kuin minun? En tiedä: täydelliselläkin teknologialla tällaisen laitteen rakentaminen vaatisi sen kalibroimista kahden subjektin verbaalisten raporttien avulla, joten kvalioiden vertailu ei ole edes periaatteessa mahdollista. (Dennett 1988, 230; 1991, 390.)

Ajatellaan sitten, että värien käännön onkin tehnyt ilkeä neurokirurgi yöllä nukkuessani. Herään ja huomaan väriaistimukseni vaihtuneiksi. Kukaan muu ei huomaa mitään merkillistä tapahtuneen, joten vian täytyy olla minussa. Nyt näyttää, että kvaliat ovat

³⁸ ”Zimbo” on mielestäni jokseenkin turha käsite – aivan kuin perinteisempään ajatukseen zombista ei joku voisi sisältyä kyky itsereflektioon.

sittenkin olemassa, niistä voi tehdä väitteitä ja niitä voi jopa testata empiirisesti. (Dennett 1988, 230; 1991, 391.) Näin ei silti ole, kuten ajatuskokeen jatkokehittely osoittaa.

Ilkeä neurokirurgi voi tehdä leikkauksen kahdella tavalla (Dennett 1988, 231). Hän joko 1) invertoi jonkin ”varhaisista” kvalioita tuottavista kanavistani, esim. näköhermossa. Tämä tietenkin vaihtaa kvaliat. Tai hän 2) invertoi muistilinkit, jotka mahdollistavat aiemmin havaittujen värisävyjen vertailun nyt havaittaviin: tämä *ei* vaihda kvalioita, vaan ainoastaan muistiin ankkuroidut taipumukseni reagoida niihin.

On waking up and finding your visual world highly anomalous, you should exclaim "Egad! *Something* has happened! Either my qualia have been inverted or my memory-linked qualia-reactions have been inverted. I wonder which!" (Dennett 1988, 231.)

Yhteys nykyisen kokemuksen ja muistin kautta tavoitetun aiemman kokemuksen välillä on ongelmallinen samalla tavoin kuin kaapeli, jonka ajateltiin liittävän yhteen kaksi subjektia. Leikkauksen kokenut *ei itse voi tietää*, turvautumatta ulkopuoliseen tietoon, kumpi näistä leikkauksista on tehty – eikä hän voi tietää itse, välittömästi, ovatko kvaliat vaihtuneet. Hänellä ei siis ole kvalioihin etuoikeutettua pääsyä (Dennett 1988, 231).

Dennett koettaa havainnollistaa invertoitujen kvalioiden absurdiutta vielä kokeella, jossa näkökenttä käännetään erityisten lasien avulla. Kuten kokeissa on havaittu, jo muutaman päivän jälkeen koehenkilöt sopeutuvat kääntämiseen hämmästyttävän hyvin. Mutta jos kysyisimme, käänsivätkö he näkökenttensä uudestaan oikein päin, vai käänsivätkö he lukemattomilla kompensatioilla ympäri kaikki reaktionsa näköaistimuksiin, he eivät osaisi sanoa. Se miltä asiat näyttävät heistä koostuu monista riippumattomista reaktiotavoista, ei yhdestä ”intrinsisesti ylösalaisesta kuvasta päässä”. (Dennett 1991, 397.) Jos vaatisimme, että heidän tulisi vastata ja että kysymykseen on oltava oikea vastaus, päätyisimme mukaelmaan invertoidun spektrin ajatuskokeesta: ”Mistä tiedän, etteivät toiset ihmiset näe kaikkea ylösalaisin (vaikka ovatkin täysin sopeutuneet siihen)”. Mutta ei ole mielekäästä väittää, että henkilön näkökentällä on ”sivuttain olemisen ominaisuus” tai ”ylösalaisin olemisen ominaisuus” riippumatta henkilön taipumuksista reagoida siihen. (Dennett 1988, 238.)

8.4.2 Argumentin premissien ja validiuden kritiikki

Zombiargumentin herättämällä kiistalla voisi helposti täyttää useammankin tutkielman, mutta nyt käsittelyssä ovat Chalmersin tietoisuuden teorian lähtökohtaiset intuitiot eivätkä niinkään näistä seuraavat argumentit. Koska zombiargumentti on kuitenkin

keskeinen Chalmersin tietoisuuden filosofialle ja osaltaan tukee väitettä vaikean ongelman olemassaolosta, referoin tässä hieman sen pätevyyyteen kohdistettua kritiikkiä.

Muotoillaan argumentti aluksi Chalmersin 2D-semantiikan kielellä³⁹ (jos jätetään tässä pois panpsykismien mahdollisuus) (Chalmers 2009, 141–153). Olkoon P täydellinen mikrofysikaalinen tieto maailmasta, ja Q jokin (mikä tahansa) fenomenaalinen väite.

- (1) $P \ \& \ \sim Q$ on kuviteltavissa (joten zombit ovat kuviteltavissa).
 - (2) Jos $P \ \& \ \sim Q$ on kuviteltavissa, niin $P \ \& \ \sim Q$ on 1-mahdollinen.
 - (3) Jos $P \ \& \ \sim Q$ on 1-mahdollinen, niin $P \ \& \ \sim Q$ on 2-mahdollinen (eli metafysisesti mahdollinen).
 - (4) Jos $P \ \& \ \sim Q$ on 2-mahdollinen, niin materialismi on epätosi.
-
- (5) Materialismi on epätosi.

Keskeinen intuitio sisältyy tietenkin premissiin (1); “One of the great battles in the Zombie Wars is over what, exactly, is meant by saying that zombies are ‘possible’.” (Yudkowsky 2008). Voidaan kysyä, ovatko zombit kuviteltavissa, ja voiko mitään päätelmiä tehdä yksilön kuvittelukyvyyn varassa (esim. P. S. Churchland 1996). Chalmers rakensi tätä vastaväitettä vastaan *ideaalisen kuviteltavuuden* käsitteen (Chalmers 2002b, 2009), jonka mukaan premississä on ajateltava ideaalisella tiedolla varustetun olennon kuvittelukykyä. Tämän mukaan väitteen S totuuden kuviteltavuus vaatii, että ei ole a priori suljettavissa pois että S on tosi (Chalmers 2006, 169).

Voidaan myös väittää, että jotta zombit olisivat kuviteltavissa, on jo ennalta oletettava vaikean ongelman olemassaolo, jonka jälkeen Chalmers käyttää zombiargumenttia vaikean ongelman tueksi (esim. Carruthers & Schier 2012; 2017, 74). Carruthersin ja Schierin (2012) mukaan Chalmersin tukeutuminen tietoargumenttiin ja Nageliin sisältää samanlaisen kehäpäätelmän. Chalmers on vieläpä myöntänyt tämän; vaikka hän onkin sanonut, että argumentit suuntautuvat materialismin kumoamiseen (Chalmers 1996, 110; 2012), hän on toisaalta myöntänyt, että ne osaltaan tukevat vaikeaa ongelmaa. Mutta tämä on hänen mielestään hyväksyttävää: ilman joitakin intuitioita filosofia ei pääsisi edes alkamaan ja tällainen toisiaan tukevien intuitioiden käyttö on filosofiassa tavallista (Chalmers 2014, 540 ja 542).

Michael Pauen (2017) esittää ajatuskokeen *osa-aikazombista*, joka on välillä fenomenaalisesti tietoinen, välillä ei. Jos zombit ovat käsitteellisesti mahdollisia, niin sitten ovat myös osa-aikazombit. Oletetaan että tällainen osa-aikazombi lyö vasaralla sormeen ei-

³⁹ Ks. luku 3.1.

kokevassa tilassa hetkellä t_1 , ja uudestaan kokevassa tilassa hetkellä t_2 , ja myöhemmin kokevassa tilassa muistelee kokemuksia hetkinä t_1 ja t_2 . Muistijäljet ovat identtiset, joten osa-aikazombin mielestä tilat ovat identtiset – joten hänen muistikuvansa ainakin toisesta näistä kahdesta fenomenalisesti huomattavasti eroavasta tilasta on pahasti väärin. Tämä on absurdia, joten siis zombit eivät ole mahdollisia.

Premissi (2) on tosi kuviteltavuuden ja 1-mahdollisuuden määritelmien perusteella. Sen sijaan (3), jonka mukaan 1-mahdollisuudesta seuraa metafyyminen mahdollisuus on kiistanalaisin; mm. B-tyyppin materialistit usein kiistävät tämän. Chalmersin ainakin aluksi näyttää onnistuvan väistämään tämän kritiikin. Chalmersin 2D-semantiikassa käsitteillä on primaari ja sekundaari intensio: väite on loogisesti mahdollinen jos se on tosi jossain maailmassa primaarin intensioon mukaan arvoituna, ja metafyyysisesti mahdollinen jos se on tosi jossain maailmassa sekundaarin intensioon mukaan arvioituna (Chalmers 1996, 68). Chalmersin mukaan kuitenkin se mikä näyttää tietoiselta kokemuksesta, *on* tietoisesta kokemuksesta ja näin primaarit ja sekundaariset intensiot yhtyvät tietoisuuden tapauksessa⁴⁰ (Chalmers 1996, 133; 2009, 149). Edellä tämä pätee väitteen Q kohdalla mutta ei välttämättä väitteen P kohdalla, jolloin argumentista seuraa toisena mahdollisuutena, että panpsykismi on totta (Chalmers 2009, 152). Toisin sanoen (3) pätee, ja se mikä on positiivisesti ideaalisesti kuviteltavissa, on myös metafyyysisesti mahdollista (tai panpsykismi on totta).

Tavallisen tulkinnan mukaan jos (4) on tosi, myös (5) on tosi (Raatikainen 2018). Voidaan kuitenkin väittää, että premisseistä (1) – (4) ei seuraa (5), koska modaalimaailma ei ole logiikan kielellä *saavutettavissa* omasta maailmastamme (Piccinini 2017). Kuviteltavuusargumenttia voi samassa muodossa soveltaa myös toimimaan dualismia vastaan (Piccinini 2017; ks. myös Balog 1999 ja Frankish 2007).

Näistä kaikista syistä zombiargumentin todistusvoima on sen prima facie -uskottavuudesta huolimatta kaukana kiistattomasta.

8.5 Itseä koskevat intuitiot

Dennettin mukaan meidän keskeinen taktiikkamme, jolla suojelemme, ohjaamme ja määrittelemme itseämme, on kertoa tarinoita. Nämä tarinat näyttävät tulevan yhdestä

⁴⁰ Panu Raatikaisen mielestä näin ei ole: argumentissa Q ei ole fenomenaalinen kvaliteetti R, vaan sellaista koskeva väitelause, eikä ole niinkään selvää, että näennäinen ja todellinen yhtyvät sen tapauksessa (Raatikainen 2018).

lähteestä: aivan kuin niiden takana olisi yksi yhtenäinen agentti. Leikkimällä ja testaamalla *mikä on minua*, fyysisesti ja mentaalisesti erottelemme, minkä havaitsemme olevan meitä itseämme ja minkä vain kuvittelemme todeksi (Dennett 1991, 428; 2017, 343–4). Kuten Anil Seth sanoo, kokemus minuudesta saattaa olla vain aivojen paras arvaus itseen liittyvien aistimusten syystä (Seth 2015, 1461; 2016). Näin rakennamme itseämme määrittelevän tarinan, joka keskittyy jonkinlaisen perusytimen, *subjektin narratiivisen painopisteen* ympärille, joka on abstraktio samaan tapaan kuin fyysikaalinen painopiste (Dennett 1991, 418 ja 429). Ja rakennamme sisäisestä minästäamme kuvan, joka toimii meille parhaiten sosiaalisina olentoina (Dennett 2015b, 409).

Normaali järjestely on ”yksi minä kehoa kohti”, mutta tämä ei mitenkään välttämätöntä: dissosiativisissa identiteettihäiriöissä näin ei näytä olevan. Samoin aivokurkiaisien halkaisuleikkauksen jälkeen molemmat aivopuoliskot osoittavat merkkejä fenomenaalista tietoisuudesta, joten jos fenomenalisuus on introspektiivinen illuusio, kummallakin puoliskolla täytyy olla kyky introspektioon ja illuusioihin (esim. Marinsek & Gazzaniga 2016). Käsitteemme tietoisuuden ykseydestä saattaa olla illuusio: tietoisuus saattaa hyvin olla osittunut (Frankish 2016b, 270).

Lisäksi, Dennett sanoo, tietoisuudella tai minällä ei tarvitse olla selvää rajaa. Koska itse ja tietoisuus ovat biologisia tuotteita, voimme odottaa, että niiden ja niiden ulkopuolisen väliset rajat ovat epäselviä. Melkein kaikki todelliset ilmiöt luonnossa kuuluvat ”samanlaisuuden ryppäisiin, joita erottaa suuret tyhjät loogisen avaruuden alueet” (Dennett 1991, 421).

Minät voisivat olla myös osittaisia, niin kuin joidenkin hyvin läheisten ihmisten väliset suhteet tuntuvat osoittavan (esim. Spurrett & Cowley 2010). Ja itse ja tietoisuus voi olla myös *aukkoinen* (Dennett 1991, 423; vrt. Blackmore 2016). Oletusta itsen olemassaolosta kritisoi jo Hume, jonka mukaan mikään ei tue jonkin ”itsen” olemassaoloa instantioituneiden kokemusten sarjan lisäksi.

Dennett toteaa, että toisin kuin yleensä oletetaan, aivot eivät ”täydennä” kuvasta (tai vastaavasti äänistä) puuttuvia osia. Silmät tekevät nopeita liikkeitä ja voivat hetkessä hakea informaatiota mistä tahansa tarkan näön alueen ulkopuolelta: antaa siis *maailman* tallentaa yksityiskohdat, siihen asti kun niitä tarvitaan. (Dennett 1991, 354–5.) Niinpä sen sijaan, kun todistetta vastakkaisestakaan ei ole, aivot hyppäävät päätelmään, että terävän näköalueen ympärillä näkyvät epäselvät läiskät ovat nekin tarkasti havaittuja. Aivot eivät

vaivaudu representoimaan jokaista yksityiskohtaa maailmasta, vaan ainoastaan *että* seinällä on maailmassa on lukemattomia yksityiskohtia. (Dennett 1991, 355.) Samaan tapaan on ajallisia reikiä: kun silmämme tekevät sakkadejaan, emme ole tietoisia puuttuvista hetkistä näkemässämme.

Dennettin mielestä näkökentän epäjatkuvuus havainnollistaa hyvin, miten tietoisuus rakentuu. Tietoisuus on *epäjatkuva*, siitä huolimatta että se näyttää ilmeisen jatkuvalta. (Dennett 1991; ks. myös Blackmore 2016). Dualistin intuitio on, että tietyt ominaisuudet kokemuksessamme ovat meille *suoraan tuttuja* – mutta miten me voisimme tietää, milloin tietomme oli läsnä koko ajan, ja milloin aivomme vain noutavat sen maailmasta? Dennett uskoo, että vain heterofenomenologinen tutkimus voi selvittää sen. (Dennett 1991, 359–60.)

Edeltävä näyttää tekevän vakavia säröjä Chalmersin päättelyyn, jossa oletetaan yksi jatkuva tietoisuus jonka sisältöjä fenomenaaliset ominaisuudet ovat. Chalmers tietenkin asettuu aivan vastakkaiselle kannalle. Chalmers on yhteisartikkelissa Timothy Baynen kanssa argumentoinut, että subjektin fenomenaalisten tilojen *joukko* on välttämättä fenomenaalisesti yhtenäinen, sillä tuskin voi kuvitella subjektilla olevan kaksi eri fenomenaalista tilaa yhtäikaa, ilman että on olemassa niitä yhdistävää fenomenologiaa (Bayne & Chalmers 2003, 36–39). Pääsy tietoisuus ja fenomenaalinen tietoisuus voivat erkaantua toisistaan, ja pääsy tietoisuuden ykseys voi kyllä hajota edellä mainituissa patologioissa mutta ei fenomenaalisen tietoisuuden ykseys. Ilmeisesti on olemassa yksi *tietoisuuden kenttä* (vrt. Searle 2004, 154 ja 295) kaikkina aikoina, ja näin ollen empirinen teesi fenomenaalisen tietoisuuden ykseyttä vastaan ei ole riittävä (Bayne & Chalmers 2003, 39). Fenomenaalisen tietoisuuden ykseyden kiistäminen tuntuu ei-kuviteltavissa olevalta, jopa epäkoherentilta (Bayne & Chalmers 2003, 55).

Esimerkiksi Blackmoren kritiikki näyttää silti torjuvan tämän käsityksen: huomiointi hyppii havainnosta toiseen ja muodostaa näistä tilapäisiä rakennelmia jotka antavat vain vaikutelman ykseydestä (Blackmore 2016, 60). Ja kuten Crick ja Koch (2003) toteavat, Baynella ja Chalmersilla kuten muillakaan tietoisuuden ykseyteen uskovilla on tuskin mitään yksityiskohtaista (ainakaan neurobiologista) aineistoa väitteidensä tueksi.

8.6 Kvaliteetti-intuitiot

quine, v. (1) To deny resolutely the existence or importance of something real or significant. "Some philosophers have quined classes, and some have even quined physical objects." Occasionally used intr., e.g., "You think I quine, sir. I assure

you I do not!" (2) n. The total aggregate sensory surface of the world; hence *quinitis*, irritation of the quine. (Dennett & Steglich-Petersen 2008.)

Dennett ei aio ryhtyä kiistelemään kvalioiden luonteesta – hänen mielestään kvalioita ei ole lainkaan olemassa, eivätkä kvalioista puhuvat edes tiedä mistä puhuvat (Dennett 1988, 229; 2005, 86). Dennettin kvalioiden kritiikki on hänen koherentein ja tehokkain hyökkäyksensä dualismia vastaan, ja sitä esittelen seuraavaksi.

Dennett luettelee neljä kvalioille yleisesti oletettua ominaisuutta. Kvaliat ovat (1) *sanoinkuvaamattomia* (ineffable): niitä ei voi kuvailla tai selittää toisille, ne voidaan tavoittaa vain suoralla kokemuksella; (2) *intrinsisiä* (intrinsic): kvaliat eivät ole relationaalisia, mikään ulkopuolinen ei määrittele niiden luonnetta; (3) *yksityisiä* (private): kaikki kvalioiden vertailut eri subjektien välillä ovat systemaattisesti mahdottomia; (4) *välittömästi käsitettäviä* (apprehensible) kokijalleen: kvalian kokeminen merkitsee että tietää kokevansa kvalian, ja lisäksi että tietää kvaliasta kaiken mitä siitä on tiedettävissä (Dennett 1988, 229; 2013, 298).

Sanoinkuvaamattomuus ja yksityisyys

Voin lukea kuvauksen esimerkiksi kalasääsken huudosta ja saada jonkinlaisen käsityksen siitä, mutta kun kuulen kalasääsken äänen luonnossa ensi kertaa, ajattelen ”Ahaa, tuolta se siis kuulostaa!” – ja ilmeisesti koen selvän, sanoinkuvaamattoman kvalian? Kuitenkin, Dennett sanoo, yhdestä kokemuksesta en voi tietää, miten yleistäisin sen muihin kalasääsken huutoihin: onko esimerkiksi puoli oktaavia matalampi huuto yhä kalasääsken huuto? Enkä voi tietää, mitkä pienet erot tai samankaltaisuudet seuraavaksi kuulemassani kalasääsken huudossa tuottaisivat täsmälleen samanlaisen kokemuksen. Enkä tiedä, olisiko minulla sama kokemus, jos myöhemmin kuulisin fysikaalisilta ominaisuuksiltaan edellisen äänen kanssa identtisen äänen. (Dennett 1988, 241.)

Dennettin mukaan, kun kuulen huudon ensi kertaa, löydän uuden ”ominaisuusilmaisimen” itsestäni. En vielä osaa kuvailla tuota ominaisuutta julkisella kielellä, mutta voin viitata siihen: se on ominaisuus, jonka havaitsin tuossa tapahtumassa. Kokemukseni kalasääsken huudosta on sanoinkuvaamaton, koska sitä ei ole vielä testattu erilaisissa olosuhteissa ja koska se on paljon informatiivisempi kuin esimerkiksi sen verbaalinen kuvaus lintukirjoissa. (Dennett 1988, 242)

Samalla tavalla punaisilta näyttävillä esineillä on objektiivisesti tuskin mitään yhteistä. Värinäkömme ei anna meille pääsyä objektien yksinkertaisiin ominaisuuksiin, vaikka

siltä näyttääkin. (Dennett 1991, 375–6.) Dennettin näkemyksen kanssa sopii yhteen esimerkiksi Derk Pereboomin (2011) analyysi, jonka mukaan näköjärjestelmä näyttää värit redusoitumattomina ja primitiivisinä kvalitatiivisinä ominaisuuksina, luultavasti koska aivojemme on helpompaa prosessoida havaintoja käyttämällä yksinkertaistettuja vakiovärejä korvaamaan värihavaintoja. Ja Pereboom ehdottaa, että introspektio näkee mielentilat samaan tapaan yksinkertaistettuina, primitiivisinä ominaisuuksina (kvalioina) (Pereboom 2011, 14–21).

Dennett vertaa Julius ja Ethel Rosenbergin tunnistuskeinoon: revitty Jell-O-laatikko toimii yksilöllisenä tarkistuksena sen vastinkappaleelle. Tässä yhdellä puoliskolla voi ajatella olevan ominaisuuden M, jonka sen vastinkappale, ”M-ilmaisim” havaitsee, eikä tätä voi mielekkäästi korvata millään monimutkaisemmalla algoritmilla. Kummallakaan puoliskolla ei ole perusteita olla olemassa ilman toista. (Dennett 1988, 242; 1991, 376). Sama pätee väreihin ja värinäköön (ja muihin aisteihin): joidenkin luonnon olioiden ”tarvitsi tulla nähdyiksi” ja toisten tarvitsi nähdä ne, joten kehittyi järjestelmä joka pyrki minimoimaan jälkimmäisten vaivan lisäämällä edellisten havaittavuutta (Dennett 1991, 377).

Tämän vuoksi kvaliat ovat niin ”sanoinkuvaamattomia”. Myös Rosenbergien Jell-O-laatikon puolikkaan muotoa ja rakennetta on lähes mahdotonta määritellä. Ainoa heti tarjolla oleva vaihtoehto on näyttää laatikon toista puolta, ja sanoa, kyseessä on muoto, joka vastaa tätä laatikon puoliskoa. Vastaavasti kun kvalioihin uskova havaitsee jonkin pinnan värin hän ei voi, kysyttäessä, minkä ominaisuuden hän havaitsee, vastata muuta kuin ”tämän”, jolloin hän uskoo osoittavansa jotakin ”sisällään” olevaa, kokemuksensa yksityistä, fenomenaalista ominaisuutta (Dennett 1991, 382–3.) Mutta tällöin hän viittaa vain ”julkiseen ominaisuuteen jolla on kartoittamaton reuna”, käyttämällä omaa henkilökohtaista, ainutlaatuista kykyään havaita tuo ominaisuus, eikä ”yksityisyys” tarkoita mitään sen enempää (Dennett 1988, 242).

Hermostollisten tapahtumien virta (esimerkiksi) näköhavainnoissa ylittää valtavasti verbaalisen raportoinnin ja muistin kyvyt. Lisäksi lukemattomat pienet eroavaisuudet yksilöiden kehityshistorioissa takaavat, että kahdella ihmisellä ei koskaan ole samanlaisia aistikokemuksia. Ne eivät siksi ole täysin kommunikoitavissa, aina tuntuu ”jotakin jäävän pois” kuvauksesta. Mutta ne ovat vain samoja funktionaalisia ominaisuuksia, ”jotka vain eivät ole vielä päässeet kuvaukseen”. (Dennett 2001a, 233.)

Kouliintumattomalle kuuntelijalle kitaran ääni on selvästi tunnistettava mutta jollakin

vaikeasti määriteltävällä tavalla. Mutta kun hän asettaa sormen otelaudalle eri kohtiin vuoron perään niin, että hän kuulee selkeämmin alkuperäisen äänen ylä-ääniä, alkuperäisen äänen ”kuvaamattomuus” purkautuu ja hän pystyy erottamaan sen osatekijöitä, jotka olivat siellä koko ajan. (Dennett 1988; ks. myös Loorits 2019, 19–20). Ennen tätä kokeilua hänellä oli käytettävissään sama informaatio ja hän pystyi hyödyntämään sitä soittaessaan, mutta ei olisi esimerkiksi kyennyt verbalisoimaan sitä (Dennett 1991, 337). Tämä muodostaa aivojen ”käyttöliittymän”, niin kuin Dennett sanoo: kvalioiden yksinkertaisuus on vain seurausta aivojen omasta rajoittuneesta pääsystä omiin monimutkaisiin prosesseihinsa (Dennett 2015b, 408). Mielenkiintoisesti Chalmersin (2018a) käsittein sanottuna Dennettin käyttäjäilluusion voi nähdä eräänä ratkaisuna tietoisuuden metaongelmaan. Chalmersin mielestä se ei kuitenkaan selitä mekanismeja, jotka synnyttävät tietoisuuden illuusion (Chalmers 2018a, 31).

Chalmersin kritiikki ei ole ihan vakuuttava, sillä Dennettin tietoisuusteoria kyllä selittää, miten illuusio syntyy ja mikä mahdollisesti on sen tarkoitus. Vaikuttaa, että Chalmers turvautuu samankaltaiseen manööveriin kuin *lisä*faktasta puhuessaan: riippumatta siitä, minkälaisen funktionaalisen selityksen illuusion synnylle annamme, jää yhä jokin selityksellinen kuilu, miksi *juuri tämä* mekanismi synnyttää illuusion. Näin määriteltynä illuusio on tuomittu pakenemaan kaikkia selitysyriytyksiä.

Intrinsisyys

Dennett kritisoi kvalioiden puolustajia kvalioiden määrittelemisestä siten etteivät kvaliat paikallistu kokemusten aiheuttamiin reaktioihin (punaisen aiheuttama hälytystila, verbaaliset raportit, vaikutukset muistiin,...) eivätkä kokemusten kausaaliisiin aiheuttajiin (aivotapahtumat), jolloin kvalioista saadaan määritelmällisellä taikatempulla ”intrinsisiä” ominaisuuksia (Dennett 1991, 386; 2001a, 233).

Of course, if you *define* qualia as *intrinsic properties* of experiences considered in isolation from all their causes and effects, and logically independent of all dispositional properties, then qualia are logically quaranteed to elude all functional analysis. (Dennett 2013, 299.)

Tämä intrinsisyys ei silti ole sen todellisempaa kuin jonkin tietyn valuutan intrinsinen arvo: jotkut naiivit amerikkalaiset tietävästi pitävät vain dollaria *oikeana* rahana, ”mutta tämä dollarin sisäinen arvo on jäävä ikuisiksi ajoiksi talousteorioiden ulottumattomiin” (Dennett 2001a, 233).

Dennett kysyy, onko parempi selittää esimerkiksi se, että yleensä inhoamme käärmeitä,

1) tietyllä ”käärme-inhotus-kvalialla” ja reaktiollamme siihen, vai 2) hermojärjestelmämme kehittyneillä taipumuksilla, jotka vaikuttavat adrenaliinin erittymiseen tai pakene-tai-taistele-rutiineihin, mikä kaikki on kokenut uudelleenkirjoittumista ja torjuntaa ja mihin kaikkeen kulttuuriset meemit⁴¹ ovat vaikuttaneet. Näistä 2) on selitys, mutta 1) vain näyttää selitykseltä: tällainen ”intrinsinen” kvalia ei selitä lainkaan subjektin reaktioita. (Dennett 1991, 385–6.)

Quining qualiassa Dennett käsittelee kukkakaalin subjektiivista makua: useimpien ihmisten mielestä keitetty kukkakaali maistuu hyvältä – Dennett taas ei voi sietää sitä. Luonnollinen päätelmä on, että *tuon maun* täytyy maistua muille erilaiselta – tai, yhtä luonnollinen päätelmä: maku on *erilainen* muille ihmisille. Aivan kuin voisi puhua siitä, millainen on ”kukkakaalin maku Dennetille hetkellä t_1 ”, ”kukkakaalin maku Dennetille hetkellä t_2 ” ja ”kukkakaalin maku minulle hetkellä t_1 ”. Mutta tällöin oletetaan, että on mahdollista erottaa kvalia kaikesta muusta mitä sen kokevalle yksilölle on tapahtumassa tai minkälaisia arvostelmia hänellä on tästä kokemuksesta, mitä siihen liittyviä muistoja ja tunteita hänellä on, miten hän on stimuloitunut muista syistä jne. Virhe on nimenomaan siinä, että edes oletetaan, että on olemassa jokin puhdistettavissa oleva, näistä kontingensseista erotettava ominaisuus (kvalia). (Dennett 1988, 228.)

Intentionaalisten järjestelmien teoria

Tilarajoitusten vuoksi mainitsen vain lyhyesti Dennettin (2009) *intentionaalisten järjestelmien teorian* (IST), jossa oletetaan organismille intentionaalisuus (eli otetaan heterofenomenologin intentionaalinen asenne) ja selitetään organismin käyttäytymistä sen avulla. Baßler (2015) täydentää onnistuneesti Dennettin teoriaa, kuvailee miten IST selittää itsemonitoroinnin kehittymisen ihmisillä ja saa lähes kaikki kvalioihin yleisesti liitetyt ominaisuudet selitettyä. (Baßler 2015, 417–21; ks. myös Dennett 2015a, 426–7.).

8.7 Esitysintuotit

Yksityinen pääsy

Jos koehenkilö tuijottaa hetken käänteisillä väreillä maalattua lippua ja siirtää sitten katseensa tyhjälle, harmaalle alueelle, jälkikuvana näkyy suunnilleen tavallisen värinen lippu (Dennett 2015b, 403–4; 2016, 71; 2017, 358.). Mutta missä, Dennett kysyy, esimerkiksi jälkikuvan ylimpänä näkyvä punainen raita sijaitsee? Sivulla ei ole punaista raitaa,

⁴¹ Ks. esim. Dennett 1991, 2013 ja 2017.

sitä ei ole myöskään hänen verkkokalvoillaan tai aivoissaan. Vaikuttaa, että hänen aivonsa ”projisoivat” sen ulkomaailmaan, ja niinpä hän tulkitsee punaisen raidan syyksi jonkin subjektiivisen ominaisuuden (kvalian). Koska hän tietää, että ulkomaailmassa ei ole mitään todellista punaista raitaa, hän on vakuuttunut, että hänen mielessään todella on olemassa ”subjektiivinen” punainen raita, *jonka olemassaolosta ei voi olla väärässä*. Hän luulee uskomuksen *intentionaalista objektia* sen syyksi. (Dennett 2017, 360 ja 363.)

Tietenkin aivoissa todella on olemassa jotakin – hermoimpulsseja, runsaasti informaatiota sisältäviä tiloja, mutta ei punaisen lipun ”kvaliaa”. Sen renderointi olisi ylimääräistä, turhaa työtä. Jos tuollainen ”punaisen raidan kvalia” todellakin olisi olemassa aivoissa, jottei se menisi hukkaan, pitäisi ilmeisesti olla jokin sisäinen tarkkailija, joka näkee sen ja toteaa ”tosiaan, tuossa on punainen raita”. Mutta se olisi jo tehdyn työn toistamista, sillä tämä tieto oli jo järjestelmässä. (Dennett 2017, 362–3.)

Erehtymättömyys

Kuten luvussa 4.4 esitettiin, Chalmers esittää niin paljon varauksia erehtymättömyysteeseensä – joka on keskeinen 1. persoonan tiedon mahdollisuudelle – että se uhkaa jäädä täysin sisällyksettömäksi. Dennett taas esittää lukuisia erehtymättömyyttä kiistäviä ajatuskokeita, joista tässä poimintoja.

Oluen makua pidetään opittuna makuna: ensimmäinen kulaus ei tietävästi kenenkään mielestä maistu hyvältä, mutta se miltä olut maistuu, muuttuu hitaasti. Jotkut sen sijaan väittävät, että olut maistuu heille täsmälleen samalta kuin ennenkin, he vain pitävät nyt *siitä* mausta. Mutta jos oluen maistaja sanoo puhuvansa siitä, miltä olut maistuu hänelle riippumatta hänen reaktiivisista taipumuksistaan, hän pettää itseään. Sillä reaktiot konstituivat kokemusta, ja niin sanotut fenomenaliset ominaisuudet osoittautuvatkin ekstrinsisiksi, relationaalisiksi ominaisuuksiksi (Dennett 1988, 236–7; 1991, 396.)

Dennettin ajatuskokeen Chase ja Sanborn ovat kaksi kahvinmaistajaa, jotka ovat olleet töissä useita vuosia Maxwell Housessa. Eräänä päivänä kahvitauolla Chase uskoutuu Sanbornille, että Maxwell Housen kahvi maistuu Chasesta samalta kuin ennenkin, mutta *hänen kykynsä arvostaa kahveja*, ts. hänen kahvimakunsa on kehittynyt niin, että hän ei enää pidä Maxwell Housen kahvista⁴². Sanborn vastaa, ettei hänkään enää pidä siitä, mutta sen sijaan – vaikka hänen kahvimakunsa on pysynyt samana – *hänen makuaistinsa*,

⁴² Chasen tilanne on vastaava kuin ajatuskokeessa, jossa Dennett saa pillerin joka poistaa hänen inhonsa kukkakaalia kohtaan mutta ei muuta kukkakaalin makua (Dennett 1988, 237).

se miten hänen makunystyränsä toimivat, on muuttunut. (Dennett 1988, 231–2.) Mitä todella on tapahtunut? Vaihtoehdot Chasen tapauksessa:

- a) Chasen kahvinmaisto-kvaliat ovat säilyneet ennallaan mutta *hänen reaktionsa niihin* ovat muuttuneet – juuri kuten hän sanoo.
- b) Chase on väärässä: hänen kvaliansa ovat muuttuneet hitaasti vuosien kuluessa, mutta hänen kahvimakunsa ei ole (hän on siinä tilassa missä Sanborn väittää olevansa).
- c) Jotakin tältä väliltä; hänen kvaliansa ovat muuttuneet jonkin verran ja hänen kahvimakunsa niin ikään jonkin verran.

Vaihtoehdot Sanbornin tapauksessa ovat, soveltaen, vastaavanlaiset. Chasen (ja Sanbornin) tilannetta voisi testata empiirisesti, esimerkiksi kokeilemalla miten hän pärjää sokkotesteissä, tai millaista hänen kahvitietoutensa on. Mutta empiiriset testit eivät riitä. Kvaliat oletettavasti vaikuttavat käyttäytymiseemme ainoastaan sikäli kuin ne vaikuttavat reaktioihimme niihin (arvostelmiimme niistä), joten mikä tahansa testi antaa suoraa todistusaineistoa vain a):n ja b):n *yhteisvaikutuksesta* (Dennett 1988, 234). Niinpä emme voi tietää kolmannen persoonan aineistosta, mikä on totuus Chasen ja Sanbornin tapauksissa. Mutta he eivät myöskään voi tietää 1. persoonan tapauksessa, onko syynä a), b) vai näiden yhteisvaikutus (Dennett 1988, 234).

Dennett kysyy: pitäisikö Chasen sanoa, että koska hän ei siedä kahvia enää, se ei enää maistu hänestä samanlaiselta? Vai sanoa, että oikeasti se kyllä maistuu yhä samanlaiselta, tai siis tavallaan, kunhan vähennetään se tosiasia, että se maistuu nyt pahalle? (Dennett 1988, 237.)

Jos pakotamme Chasen valitsemaan jommankumman näistä vaihtoehdoista, hän on pulassa. Hän voi sanoa, että näin jotenkin pakotamme erilleen hänen esiteoreettisen kvalian käsitteensä piirteet, jotka aiemmin olivat luontevasti yhdistyneinä – mutta näin hän vain tukee Dennettin väitettä, että tavallisessa arkipsykologiassa ei ole kunnollista perustaa kvalian käsitteelle (Dennett 1988, 237). Hänen intuitiiviset arvostelmansa hänen kvalioidensa samankaltaisuudesta eivät ole sen luotettavampia kuin hänen arvostelmansa esimerkiksi siitä, onko hänellä kuumetta vai onko hänellä vain kuumeinen olo. Korkeintaan hän voi sanoa, että ”relaatio hänen kahvin maistamisensa ja sen arvostelun välillä on muuttunut”, mutta tällöinkään hänen käsityksensä hänen hermostollisista tapahtumistaan eivät riipu pelkästään ”välittömistä” vaikutelmista vaan myös

suhteesta muistoihin aiemmista vaikutelmista.

Arkipsykologian mukaan kvaliat ovat ominaisuuksia, joiden olemassaolosta tai luonteesta ei voi erehtyä – vieläpä niin, että kun koemme kvalian, voimme tietää kirjaimellisesti täydellisesti millaisia ne ovat: kvalia Q on aina mitä se on eikä mitään muuta, sitä voi käyttää vertailussa, se joko on tai ei ole, jne. (Lewis 1995, 141–2). Tällöin, jos lisäksi Chalmersin teoria fenomenalisista arvostelmista hyväksytään, myös fenomenaliset arvostelmat voivat olla erehtymättömiä. Mutta Dennettin edeltävän kritiikin perusteella väite ei näytä pitävältä. Kvalioita on pidettävä ennemminkin kvalia-arvostelmista johdettuina *loogisina konstruktioina*: subjektilla on kvalia F, jos ja vain jos hänellä on arvostelma, että hänen kokemuksessaan on kvalia F. Näitä arvostelmia on siten pidettävä *konstituivina tekoina*, jotka synnyttävät kvaliat samaan tapaan kuin kirjailijat määrittelevät esimerkiksi henkilöhahmojensa ulkonäön. Kvaliat ovat siis samantapaisia konstruktioita kuin novellin henkilöt, ja subjektilla on vain tässä mielessä ehdoton oikeus sanoa, erehtymättömästi, millaisia ne ovat. (Dennett 1988, 233.)

Epätavalliset näköhäiriöt

Lisätukea väitteilleen Dennett hakee epätavallisista näköhäiriöistä (puolet näkökenttää on harmaasävyissä, kyvyttömyys yhdistää värit niiden nimiin ja tietämättömyys tästä puutteesta, jne.): jos tutkittava henkilö ei valita näköhäiriöstä mutta tutkimukset osoittavat sellaisen, ovatko hänen kvaliansa muuttuneet vai eivät? (Dennett 1988, 238.)

His capacities to discriminate are terribly impaired, but luckily for him, his inner life is untouched by this merely public loss! (Dennett 1988, 238.)

Dennettin mukaan ei ole mielekästä yrittää löytää vastauksia kysymyksiin kvalioista, joita tällaiset erikoiset vaivat herättävät (vrt. Searle 2004, 288). Mitkä vain vaikutukset hermostossa voivat olla irtautuneita luonteenomaisista yhteyksistään, eikä intuitiomme kerro meille mitkä näistä ovat olennaisia kvalioiden identiteetille tai pysyvyydelle. Osoittautuu vain, että arkipsykologian mukainen käsityksemme kvalioista joutuu ”täydelliseen sekasortoon näiden esimerkkien edessä”. (Dennett 1988, 240.)

8.8 Arvointuutiot

Chalmersin päättely ei nojaa arvointuutioihin, mutta niiden olemassaolon voi nähdä tueksi illusionismille. Koska tietoisuudella on niin vahva vaikutus käyttäytymiseemme, sen olemassaololle voi antaa evolutionaarisen selityksen: se esimerkiksi antaa vaikutelman

vahvasta sisäisestä elämästä, mikä antaa kokijalleen syvemmän kiinnostuksen olemassa-oloon ja toimintaan ympäristössään (Frankish 2016a, 29 ja 37).

Ei kuitenkaan pidä olettaa kokemuksellista tietoisuutta todella olemassa olevaksi, Dennett sanoo ja vertaa kysymykseen ”mitä adaptiivista etua terveydestä on?” Epäilemättä on oltava terve, jotta voisi kiipeillä vuorilla, ja voimistelijan pitää olla terve, mutta näistä ei pidä päätellä jotakin mystistä ”terveyttä”, joka näillä suorittajilla on heidän kyvykkyyksiensä lisäksi. Niin kuin ”terveys” olisi jokin *lisäilmiö* jonka voi lisätä osien toimintaan tai vähentää niistä. (Dennett 1995, 324–5.)

8.9 Kattavuusintuities

Kun Chalmers on esitellyt lähtökohtansa ja premissinsä hän toki esittää väitteitä tietoisuuden esiintymisestä eri järjestelmissä, mutta kattavuusintuutioilla – esimerkiksi intuitiolla, jonka mukaan tietoisuus voi nousta vain biologisista aivoista kuten Searle (1980, 2004) esittää – ei ole roolia argumentoinnissa. Tietysti vaikkapa organisatorisen invarianssin periaate on intuitioon perustuva väite, mutta sen voinee tämän tutkielman kontekstissa sivuuttaa, kun huomaamme Chalmersin kaikkein ongelmallisimpien intuitioiden sijaitsevan jo hänen tietoisuuden teoriansa alkuoletuksissa.

9 Yhteenveto Dennettin kritiikistä

Miten Dennett onnistuu? Kuten aiemmin todettiin, Dennett ei useinkaan koeta etsiä virheitä Chalmersin argumentoinnista, jonka perinpohjaisuuteen Dennett tuntuu olevan suorastaan turhautunut⁴³:

Chalmers never *leaps* to conclusions; he *oozes* to conclusions, checking off all the caveats and pitfalls and possible sources of error along the way with exemplary caution [...]. (Dennett 2012, 90.)

Sen sijaan hän pyrkii osoittamaan, että Chalmersin teoriarakennelman alkuoletukset perustuvat sekaviin arkipsykologisiin käsityksiin, minkä jälkeen Dennett esittää näille käsityksille empiiriseen tieteeseen perustuvat selitykset. Hän ei monestikaan onnistu varsinaisesti *todistamaan* Chalmersin olevan väärässä, paitsi niiltä osin kun Dennett onnistuu osoittamaan arkipsykologisten käsitteiden ristiriitaisuudet. Dennettiä tällainen

⁴³ Hän ei ole ainoa. Yudkowsky (2008): ”Chalmers is one of the most frustrating philosophers I know. — Chalmers does this really *sharp* analysis... and then turns left at the last minute. He lays out everything that's wrong with the Zombie World scenario, and then, having reduced the whole argument to smithereens, calmly accepts it.”

todistelu ei edes kiinnosta; hän lainaa Donald Hebbiä: ”If it isn’t worth doing, it isn’t worth doing well” (Dennett 2006).

Dennettin lähestymistavan ongelma näyttää olevan täydellinen ensimmäisen persoonan ja sitä kautta intersubjektiviivisen tiedon arvon kiistäminen – vaikka vaikuttaa, että varsinkin tietoisuuden tapauksessa kolmannen persoonan menetelmiä on täydennettävä ensimmäisen persoonan menetelmillä (Velmans 2007, 723). Miksei minun tietoinen kokemukseni voisi olla *osa* havaintoaineistoani, Levine kysyy: ilman sitä *minun episteeminen asemani* jäisi vajavaisesti esitetyksi (Levine 1994, 118). Tästä huolimatta näyttää, että Dennett saa purettua auki monia realistien intuitioita.

Metafyysiset intuitiot

Dennett pitää ilmeisenä, että usko vaikean ongelman olemassaoloon on perusteeton. Hän ei voi todistaa sitä, mutta esittää vakuuttavan selityksen illuusion synnylle: monimutkaisuudesta, taitavuudesta ilman ymmärrystä ja bayesilaisista ennustuksista syntyy introspektiivisesti läpinäkymätön tietoisuus, johon näyttää liittyvän ”vaikea ongelma”. Lisäksi on olemassa empiirisiä todisteita, että vaikea ongelma ei ole niin ilmeinen kuin Chalmers väittää.

Selittävät intuitiot

Dennett myöntää intuitioiden olemassaolon mutta ei hyväksy niille mitään episteemistä erityisarvoa. Dennettin mukaan Chalmers olettaa sen mikä pitäisi selittää – ja Dennettin tietoisuusteoria tarjoaa selityksen näennäisille lisäfaktoille. Dennettin teoria on taloudellisempi: Chalmers joutuu postuloimaan ylimääräisiä fundamentaalisia ominaisuuksia ja luonnonlakeja, kun taas Dennett selviää fysikalistisilla käsitteillä ja ominaisuuksilla.

Tietointuitiot

Dennett kieltää täysin 1. persoonan tiedon tieteellisen merkityksen. Argumentointi perustuu a) väitteeseen, että tieteellinen tieto voi perustua vain 3. persoonan aineistoon, b) fenomenaalisten käsitteiden sisäiseen ristiriitaisuuteen ja c) 1. persoonan tiedon epäluotettavuuteen.

Mikäli a) hyväksytään, joudutaan hylkäämään pitkä perinne yritykseen hankkia ensimmäisen persoonan tietoa, mutta vastaavasti a):n kieltäminen hylkää empiirisen tieteen keskeisen periaatteen. Väite on silti hyvin epäintuitiivinen, ja näin jyrkkänä vaatii hylkäämään mahdollisuuden kokemukselliseen tietoon tavalla, jota on vaikea hyväksyä. Sen sijaan Dennettin b) ja c) –argumentit ovat vakuuttavia: fenomenaalisisilla käsitteillä ei

ole selvästi määriteltyä vastinetta todellisuudessa ja subjektiivinen tieto on erittäin epäluotettavaa.

Dennettin argumentit Jacksonin Mary-argumenttia kohtaan jäävät heikoiksi. Myöskään Dennettillä ei ole esittää oikeastaan muita perusteluita epifenomenalismia vastaan, kuin että epifenomenalismi vaikuttaa hänen intuihionsa mukaan väärältä. Tämä pitävän argumentin puute on samankaltainen kuin Dennettin kyvyttömyydessä todistaa, että vaikeaa ongelmaa ei ole.

Modaali-intuihio

Dennett rakentaa hyvät perustelut sille, että monimutkaisuudesta voi syntyä jotakin tietoisuuden kaltaista, josta edelleen seuraa, että zombit (niin monimutkaisina kuin ne pitäisi kuvitella) eivät välttämättä ole sen enempää kuviteltavissa kuin mahdollisiakaan. Yhdessä zombiargumentin muiden ongelmien kanssa zombiargumentin voima tukea Chalmersin tietoisuusteoriaa on hyvin kyseenalainen. Myös realistien esittämä invertoitujen kvalioiden argumentti näyttää kaatuvan kvalian käsitteen epämääräisyyteen. Kuviteltavuusargumentista Chalmers ei siten näytä saavan riittävästi tukea teorialleen.

Itseä koskevat intuitio

Dennettin empiiriseen tutkimukseen tukeutuvat argumentit osoittavat vakuuttavasti, että tietoisuuden atomisuus ja jatkuvuus ovat kestäättömiä oletuksia. Tämä voisi sopia yhteen Chalmersin panpsykististen spekulointien kanssa – mutta Chalmersin tietoisuuden teoria rakentuu siltikin vahvasti oletukselle yhdestä subjektista. Mary on se, joka kokee ja vertailee kokemustaan aikaisempaan. On jonkinlaista olla *minä*. Ja kokemuksellisen tietoisuuden 1. persoona on yksikössä.

Kvaliteetti-intuihio ja esitysintuihio

Dennettin kvaliakritiikki on niin perinpohjaista, että sen jälkeen kvalioita ei voi pitää enää mielekkäinä käsitteinä tai edes minään todella olemassa olevina ominaisuuksina.

Arvointintuihio ja kattavuusintuihio

Dennettin arvointintuihioihin liittyvät ajatukset ainoastaan *sopivat yhteen* Dennettin tietoisuusteorian kanssa mutta eivät tee työtä Chalmersin näkemyksen kumoamiseksi. Kattavuusintuihio ovat, kuten aiemmin todettu, epärelevantteja tässä yhteydessä.

Yhteenveto

Mitä edeltävän perusteella Chalmersin teoriaa tukemaan jää? 1) Vaikean ongelman olemassaoloa ei voi todistaa mutta sen sijaan on empiirisiä todisteita siitä, että se ei ole niin ilmeinen explanandum kuin realistit väittävät. Lisäksi, Chalmersin erehtymättömyysteisiin tekemien varausten ja Dennettin kritiikin perusteella, fenomenaalista tietoisuutta koskevia arvosteliamme tuskin voi väittää erehtymättömiksi. 2) Selitysargumentti pysyy hyvin heikkona pystyssä, jos hyväksytään, että ihmiset todella voivat tuntea kipua. Kivun kokemus on silti selitettävissä illusionistisesti. 3) Zombiargumentti näyttää kaatuvan prima facie -uskottavuudestaan huolimatta (jonka suhteen intuitiot vaihtelevat!) useista kohdin päättelyketjua – mutta ennen kaikkea Dennettin selitys tietoisuuden syntymisestä (monimutkaisuudesta) on elegantimpi eikä tuo niin paljon metafysisistä painolastia mukanaan kuin Chalmersin selitys. 4) Kvalioiden inversio -argumentti taas toimii vain, jos kvalioilla on jokin vastine todellisuudessa, mitä niillä ei Dennettin kritiikin perusteella selvästikään ole. 5) Episteemisen epäsymmetrian argumentti ei sekään Dennettin hyökkäyksen alla pysty todistamaan oikeastaan muuta kuin että meillä *joko on tai näyttää olevan* omakohtaista kokemusta tietoisuudesta.

Jää 6) tietoargumentti. Mutta vaikka Dennett ei onnistu sitä kumoamaan, argumentti on kaikkea muuta kuin kiistaton (yleisimmät vastaväitteet ks. esim. Torin Alter 2007, 402–3). Ja vaikka argumentti menisikin läpi, mitä se oikeastaan onnistuu todistamaan? Että meillä on jokin jäännös, jokin kokemukselliseksi sanottava hiven psykologiassamme *jota ei vielä ainakaan ole selitetty* fysikaalisesti. Näin muotoiltuna väite sisältää tietenkin myös intuitiivisen käsityksen, että kaikki on fysikaalista ja tietoisuus sekin on jollakin tavalla fysikaalisesti selitettävissä (esim. Chalmers 1996, 245), mutta tämän käsityksen tukena on myös tieteellisen teorianmuodostuksen periaatteet kuten säästäväisyys ja yksinkertaisuus. Vaihtoehtona olisi vain realistien intuitiivinen käsitys, josta edelleen Chalmersin argumentointi näyttäisi johtavan päätelmän, että kokemuksellisuus olisi postuloitava uutena fundamentaalisen ominaisuutena. Se on kuitenkin varsin rohkea loikka ilman tähän asti verrattoman menestyksekkään empiirisen tieteen tukea.⁴⁴ Näin ollen ei näytä jäävän mitään pakottavaa syytä hyväksyä Chalmersin ominaisuusdualismia tietoisuuden teoriaksi.

⁴⁴ On tosin myös ajateltavissa, että ominaisuusdualismi on *vähemmän* metafysisinen kanta kuin materialismi: materialismi perustuu oletukseen, jonka mukaan kaikki oleva supervenioi metafysisellä välttämättömyydellä fysikaalisen yllä, mikä on vahva metafysisinen oletus (esim. Raatikainen 2018).

10 Yhteisen maaperän etsintää

Carruthers ja Schier (2017) kiinnittävät huomiota kohtaan *The Conscious Mindin* alussa, joka antaa vaikutelman, että Chalmers lähti vain tutkimaan erään idean seurauksia, ilman varsinaista pyrkimystä vakuuttaa toisin ajattelevia (Carruthers & Schier 2017, 71).

Chalmers kirjoittaa:

This book may be of intellectual interest to those who think there is not much of a problem, but it is really intended for those who feel the problem in their bones. By now, we have a fairly good idea of the sort of theory we get if we assume there is no problem. In this work, I have tried to explore what follows given that there is a problem. The real argument of the book is that if one takes consciousness seriously, the position I lay out is where one should end up. (Chalmers 1996, xiii.)

Näyttää todellakin siltä, että alun perin puhtaassa matematiikassa koulutuksensa saanut Chalmers päätti soveltaa myös filosofiassa aiemman oppialansa apriorista metodologiaa ja seurasi, mihin tuloksiin hänen alkuoletuksensa loogisesti johtaisivat. Hänen argumentointinsa onkin tarkkaa, erityisesti verrattuna Dennettin aivan erilaiseen etenemistapaan. Vaikean ongelman ”vaikeus” premissinä johti tavallaan luonnollisesti *The Conscious Mindin* negatiiviseen painotukseen, jonka mukaan kaikki reduktiiviset selitykset (ja näköjään kaikki selitykset lukuun ottamatta Chalmersin erittäin spekulatiivisia yrityksiä) välttämättä epäonnistuisivat. Positiivisissa selitysy yrityksissään Chalmers sittemmin joutui ottamaan tukea useammastakin tärkeästä arkipsykologisesta intuitiosta. Lisäksi, vaikka Chalmers väitti, että hän ei yrittänyt todistaa esimerkiksi (intuitioiden varassa seisovien) kuviteltavuus- ja tietargumentin avulla vaikean ongelman olemassaoloa (esim. Chalmers 2012), hän näki niiden silti rakentavan vaikeaan ongelmaan liittyvää intuitiota. Niinpä hän pystytti nämä eri intuitiot tukemaan toisiaan ja sitten perusteli, että tällaisen dialektiikan varassa filosofia usein lepää (Chalmers 2014, 540).

Ehkä tämä parhaimmillaan on ongelmanratkaisua edistävää ”hyvänlaatuista” kehämäisyyttä niin kuin Jennifer Matey (2012) sanoo, mutta toisaalta näyttää, että Chalmersin metodi käy myös esimerkiksi eräästä analyyttisen filosofian ongelmasta. Kuten Aaron Preston (2015) kirjoittaa, nykyisen analyyttisen filosofian skandaalimainen metodologinen puute on sen *luonteenomainen* taipumus tukeutua – tai kuten Hintikka (1999, 130) sanoo: sokea usko – intuition, vailla konsensusta sovellettavan intuition luonteesta tai siitä, miten se voisi toimia (Preston 2015, 41–42).

Myöskin monet Chalmersin uudemmat siirrot kuten turvautuminen toisiinsa tukeutuviin intuitioihin, Mooren inspiroimat todistukset tai spekulointi realisationismilla antavat

vaikutelman kuin vaikean ongelman alkuperäinen liikevoima alkaisi jo uhkaavasti hiipua, mikä saa myös Chalmersin tietoisuuden teorian ikävästi vaikuttamaan juuri sellaiselta filosofiselta harharetkeltä, jollaisista Dennett (2006) varoitti nuoria tutkijanalkuja. Dennett esitteli shakista pienellä sääntömuutoksella luomansa chess-pelin, joka mitättömästä merkityksestään huolimatta on mainiosti kykenevä generoimaan loputtomasti syvällisiä väittelyitä, nokkelia vasta-argumentteja ja pikkutarkkoja analyyseja.

The Meta-problem of consciousness etsii ratkaisua ”tietoisuuden metaongelmaan” ja nimenomaan aiheutraalia ratkaisua, joka ei ottaisi kantaa realismiin tai illusionismiin puolesta, mutta on ilmeistä, että Chalmers ei enää ehdottomasti kannata aiempaa dualismiaan. Ehkä tietoisuus ei sittenkään ole lisä fakta: “Speaking for myself, I am not sure whether a further-fact view or a deflationary view is correct.” (Chalmers 2010, 53). Tietoisuudella ei ehkä sittenkään ole ykseyttä: “We will not aim to conclusively prove the unity thesis [of consciousness] in this paper, and indeed we are not certain that it is true.” (Bayne & Chalmers 2003, 24.) Hän näyttää myös hyväksyneen kvalia-käsitteen hylättäväksi: ”However, the qualia view is widely rejected these days, even as an account of how experiences seem to us introspectively.” (Chalmers 2018a, 27). Ja niin edelleen.

Chalmersin viimeisin pyrkimys vaikuttaa olevan Dennettin heterofenomenologian jos ei kirjaimen niin ainakin hengen mukaiselta. Ja kun Chalmers pohtii ensimmäisen persoonan tutkimuksen metodien ja raporttien formalismin puutteellisuutta, Dennettin on helppo tulkita Chalmersin pyrkimykset juuri siksi mitä Dennett on tehnyt:

What a good idea: we can let subjects speak for themselves, in the first-person, and then we can take what they say seriously and try to systematize it, to capture the structure of their experience! And we could call it heterophenomenology. (Dennett 2001b.)

Chalmers pitää hyödyllisenä selityksenä, että ongelmaintuitioihin liittyvät mekanismit voivat johtua joistakin hyödyllisistä valinnoista kognitiivisessa järjestelmässä: esimerkiksi kuoleman pelko ja itsesäilytysvaisto ovat epäilemättä hyödyllisiä organismille. Chalmers hyväksyy myös mahdolliseksi kulttuuriset vaikutteet, kuten että olemme kaikki liikaa Descartesin vaikutuksen alaisia. (Chalmers 2018a, 33.) Meemeistä Chalmers ei tietenkään puhu, mutta hänen näkemyksensä alkavat tässä kovasti muistuttaa Dennettin esittämiä.

Chalmers pitää myös kiinnostavana Pereboomin (2011, 2016) ajatusta, jonka mukaan kognitiivisella järjestelmällä ei ole pääsyä mekanismeihin jotka synnyttävät mielentilan,

joten järjestelmä pitää mielentilaa välittömästi annettuna (Chalmers 2018a, 25). Chalmers pohdiskeli tätä jo 1996 (s. 289–93), ja vaikka hän käsitteli fenomenaalisten arvostelmien syntyä, näkemyksestä on vain lyhyt askel Dennettin käsitykseen siitä, miten ”kvalitatiiviset ominaisuudet” rakentuvat.

Andy Clarkin mielestä Chalmers oli tuolloin hyvin lähellä vaikean ongelman hajottamista (A. Clark 2000, 31–32). Chalmers analysoi introspektiivisen läpinäkymättömyyden merkitystä: koska järjestelmällä ei ole pääsyä prosesseihin jotka tuottavat arvostelman ”tuossa on punainen kuppi pöydällä”, on luonnollista että järjestelmä muodostaa käsityksen että havaitut erot ovat raa’asti ja primitiivisesti erilaisia, vain kvaliteettiensa määrittämiä – ja nyt tämä primitiivinen suhde näyttää arvoitukselliselta ja ”kokemuksen” ja ”kvalian” käsitteet seuraavat luonnollisesti perässä (Chalmers 2018a, 25–28). Mallin ongelmana Chalmersin mukaan on, että se sopii kuvaamaan myös tietoisuuteen liittyviä uskomuksia ja käsitteitä, jotka eivät kuitenkaan synnytä ongelmaintuitioita niin vahvasti kuin fenomenaaliset kokemuksemme. Chalmers pohtii ratkaisuksi suoraa *tuttuusrelaatiota*, joka meillä näyttää olevan tietoisuuden laadullisiin ominaisuuksiin mutta ei niitä koskeviin uskomuksiin (Chalmers 2018a, 29). Tämän tuttuuden lähteenä on kuitenkin introspektiivinen läpinäkymättömyys. Chalmers esittää lupaavimmaksi lähestymistavaksi jotakin mitä Dennett voisi sanoa:

(M) ”We have introspective models deploying introspective concepts of our internal states that are largely independent of our physical concepts. These concepts are introspectively opaque, not revealing any of the underlying physical or computational mechanisms. Our perceptual models perceptually attribute primitive perceptual qualities to the world, and our introspective models attribute primitive mental relations to those qualities. We seem to have immediate knowledge that we stand in these primitive mental relations to primitive qualities, and we have the sense of being acquainted with them.” (Chalmers 2018a, 34.)

Chalmersin mukaan kuvaus on neutraali: ”A realist can take it as a picture of how genuine consciousness seems to us introspectively. An illusionist can take it as a picture of the illusion of consciousness.” (Chalmers 2018a, 34.) Chalmers sanoo silti *suosivansa* realistista tietoisuuden teoriaa: ”our experiences seem to primitively acquaint us with qualities in the environment, and these experiences are themselves objects of acquaintance” (Chalmers 2018a, 39). Chalmers ei siis myönnä, että olisi kääntymässä illusionismiin, mutta sanoo, että jos metaongelma halutaan sijoittaa vanhan reduktionistien ja ei-reduktionistien välisen kiistan maastoon, sen voi nähdä ”ystävällisenä kädenojennuksena

vastapuolelle” (Chalmers 2018a, alaviite s. 9).

Kuitenkin, jos (M) on Chalmersin kanta, niin introspektiivisen läpinäkymättömyyden periaatetta ja ennakoivan prosessoinnin teoriaa voisi ajatella silloittamaan kuilua realistien ja illusionistien välillä (esim. Havlík ym. 2017; Hobson & Friston 2014). Kiista ei varmaankaan ratkea pelkästään filosofisella argumentoinnilla, empiirinen tutkimus on tässä myös merkittävän tärkeässä asemassa.

Ennakoivan prosessoinnin teoriasta tuskin tietenkään saadaan tietoisuuden ”kaiken teoriaa”; ennakointi ja havainnointi toimivat yhdessä vaihtelevissa määrin, ja havaitsemisen ja kognition raja on epäselvä (A. Clark 2013, 190). Esimerkiksi Buckner (2018) on (erityisesti Locken empiristisen mielen filosofian inspiroimana) tutkinut keinotekoisia hermoverkkoja ja selittänyt miten menestyksekkäät järjestelmät abstrahoivat syötteestä rikkaasti strukturoituja abstraktioita, joissa yksittäiset havainnot säilyvät ”tallennettuina”. Nisäkkäillä rakenne on ilmeisesti, vaikkakin monimutkaisempi, samankaltainen (ks. esim. P. M. Churchland 1989, 197–230). Nämä abstrahoinnit ovat läpinäkymättömiä introspektiolle (Buckner 2018, 5348). Kvalioihin sovellettuna – jos käsitettä Dennettin kritiikin jälkeen vielä sopii käyttää – tämä malli voisi onnistuneesti kuvata kvalioiden samanaikaista sanoinkuvaamattomuutta, näennäistä atomisuutta ja sisäsyntyisyyttä, kohtuullisen vähäisellä jännitteellä Chalmersin nykynäkemyksiin ja sopusoinnussa Dennettin pandemoniumin kanssa.

Bucknerin teoria lähestyy mielenkiintoisella tavalla Crickin ja Kochin (2003; ks. myös Loorits 2019, 14–15) neurobiologista teoriaa: kvaliat realisoituvat pienten neuroniryhmien yhteyksissä, niin sanotuissa *essentiaalisissa noodeissa*, monimutkaisena, holistisena struktuurina joka koostuu lukemattomista *henkilökohtaisista* assosiaatioista, esimerkiksi subjektin keltainen kvalia subjektilla yhdistyy hänen maistamaansa sitruunaan, van Goghin maalaukseen jonka hän on nähnyt jne. Näiden teorioiden lähempään tarkasteluun ei tässä tutkielmassa ole valitettavasti tilaisuutta.

Subjektiiivisuus vaikuttaa silti joltakin joka ei ole reduktiivisesti selitettävissä, mutta nähdäkseni ei-reduktiivinen jäänne on vain *indeksisyys*. Kuten Dennett argumentoi, ”subjektiiivinen näkökulma” hajoo kun sitä tarkastellaan lähemmin (esim. Dennett 1991, luku 5). Kyse on vain raa’asta tosiasista (vrt. ”miksi tämä pöytä on juuri tämä pöytä eikä tuo pöytä tuolla?”), joka ei anna mitään syytä kiirehtiä muokkaamaan metafysiikkaa. Myös Patricia Churchlandin kritiikki vaikuttaa yhä pätevältä: vaikean ongelman vaikeus

ei ole tätä ongelmaa koskeva tosiasia sen enempää kuin maailman metafyyminen ominaisuuskaan, vaan ainoastaan epistemologinen *meitä* koskeva tosiasia (P. S. Churchland 1996, 406).

Lisäksi ”subjektiivisuus” on epätarkasti määritelty käsite – kuten muutkin vaikean ongelman yhteydessä käytettävät käsitteet (Mandik 2016). Niihin liittyy kehämäisyyttä niin kuin Chalmersin intuitioihinkin ja niiden ajateltu merkitys vaihtelee huomaamatta käyttäjän mukaan. ”Millaista-on-olla” on epätarkka ja monitulkintainen ilmaus (Farrell 2016), ”kvalia” on Dennettin kritiikin perusteella vain arkipsykologinen nimilappu, ”fenomenaalisella ominaisuudella”, vaikka se onkin tekninen käsite, ei ole tarkkaa määritelmää (Dennett 2005, 79). ”Kipu” on epäjohdonmukaisesti käytetty käsite (Dennett 1978), ja ”tietoisuus” taas ei vastaa mitään yhtä ilmiötä todellisuudessa ja on siten käyttökelvoton käsite tieteelliseen tutkimukseen (Blackmore 2016; Irvine 2017).

Käytettävien käsitteiden epämääräisyys heijastelee vain sitä tosiseikkaa, että tutkittava kohde ei ole niin selvästi kiinnitettävissä kuin esiteoreettiset oletuksemme antavat aiheita uskoa. Tietoisuuden perustana olevat prosessit ovat monimutkaisia ja erityisesti keholliisuuden rooli voi olla hyvinkin suuri. Kuten Elizabeth Irvine kirjoittaa:

The contribution [of this paper] is instead is [!] showing what actually seems to happen when you go out and ‘do the science’, and that it is perhaps not very palatable to either neo-dualists, or materialists looking for a simple solution. ‘Explaining consciousness’ will be spectacularly messy [...]. (Irvine 2017, 105, virhe lähteessä.)

Selitys tulee epäilemättä olemaan sotkuinen, ja uskoakseni lähempänä Dennettin monella rintamalla tapahtuvaa ja moniäänistä kuvausta kuin Chalmersin johdonmukaista järjestelmän rakentamista apriorisista lähtökohdista käsin. Chalmersin puheenvuorot ovat joka tapauksessa piristäneet reippaasti mielenfilosofista keskustelua: Dennett ja muut realismin kriitikot ovat joutuneet tosissaan paiskimaan töitä vastatakseen Chalmersin esittämään haasteeseen.

Lähteet

- Alter, T. 2007. The knowledge argument. Teoksessa M. Velmans & S. Schneider (toim.), *The Blackwell Companion to Consciousness*. Blackwell Publishing, 396–405.
- Balog, K. 1999. Conceivability, possibility, and the mind-body Problem. *The Philosophical Review*, 108(4), 497–528.
- Balog, K. 2016. Illusionism’s discontent. *Journal of Consciousness Studies*, 23(11–12), 40–51.
- Baßler, D. H. 2015. Qualia explained away. Teoksessa T. Metzinger & J. M. Windt (toim.), *Open MIND* [online]. MIND Group, 413–423. Noudettu 15.12.2018 osoitteesta <http://cognet.mit.edu/pdfviewer/book/9780262340281/chap30>
- Bayne, T. J., & Chalmers, D. J. 2003. What is the unity of consciousness? Teoksessa A. Cleeremans (toim.), *The Unity of Consciousness: Binding, integration, and dissociation*, 23–58.
- Blackmore, S. 2016. Delusions of consciousness. *Journal of Consciousness Studies*, 23(11–12), 52–64.
- Block, N. 1978. Troubles with functionalism. *Minnesota Studies in the Philosophy of Science*, 9, 261–325.
- Block, N. 1995. On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(02), 227–247.
- Bourget, D., & Chalmers, D. J. 2014. What do philosophers believe? *Philosophical Studies*, 170(3), 465–500. <https://doi.org/10.1007/s11098-013-0259-7>
- Broackes, J. 2007. Black and white and the inverted spectrum. *The Philosophical Quarterly*, 57(227), 161–175.
- Buckner, C. 2018. Empiricism without magic: transformational abstraction in deep convolutional neural networks. *Synthese*, 195(12), 5339–5372.
- Campbell, K. 1970. *Body and Mind*. Macmillan.
- Carruthers, G., & Schier, E. 2012. Dissolving the hard problem of consciousness. 4th Online Consciousness Conference. Noudettu 20.2.2019 osoitteesta <https://consciousnessonline.wordpress.com/2012/02/17/dissolving-the-hard-problem-of-consciousness/>
- Carruthers, G., & Schier, E. 2017. Why are we still being hornswoggled? Dissolving the hard problem of consciousness. *Topoi*, 36(1), 67–79.

- Chalmers, D. J. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York ; Oxford: Oxford University Press.
- Chalmers, D. J. 1997. Moving forward on the problem of consciousness. *Journal of Consciousness Studies*, 4(1), 3–46.
- Chalmers, D. J. 1998. On the search for the neural correlate of consciousness. Teoksessa S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (toim.), *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*. MIT Press, 219–230.
- Chalmers, D. J. 2002a. Consciousness and its place in nature. Teoksessa D. J. Chalmers (toim.), *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press, 247–272.
- Chalmers, D. J. 2002b. Does conceivability entail possibility? Teoksessa T. S. Gendler & J. Hawthorne (toim.), *Conceivability and Possibility*. Oxford University Press, 145–200.
- Chalmers, D. J. 2003. The content and epistemology of phenomenal belief. Teoksessa Q. Smith & A. Jokic (toim.), *Consciousness: New Philosophical Perspectives*. Oxford University Press, 220–272.
- Chalmers, D. J. 2004a. How can we construct a science of consciousness? Teoksessa M. S. Gazzaniga (toim.), *The Cognitive Neurosciences III*. MIT Press, 1111–1119.
- Chalmers, D. J. 2004b. Phenomenal concepts and the knowledge argument. Teoksessa P. Ludlow, Y. Nagasawa, & D. Stoljar (toim.), *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. MIT Press, 269–298.
- Chalmers, D. J. 2006. Phenomenal concepts and the explanatory gap. Teoksessa T. Alter & S. Walter (toim.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford University Press, 167–194.
- Chalmers, D. J. 2009. The two-dimensional argument against materialism. Teoksessa D. J. Chalmers (toim.), *The Character of Consciousness*, 141–205. Ensimmäisen kerran julkaistu suppeampana teoksessa McLaughlin, B. & Beckermann, A. & Walter, S. (toim.) *Oxford Handbook of Philosophy of Mind*. Oxford University Press.

- Chalmers, D. J. 2010. The singularity: A philosophical analysis. *Journal of Consciousness Studies*, 17(9–10), 7–65.
- Chalmers, D. J. 2012. [ei otsikko]. David Chalmersin kommentti Carruthersin ja Schierin artikkeliin Dissolving the hard problem of consciousness. 4th Online Consciousness conference. 20.2.2012 02:05. Noudettu 20.2.2019 osoitteesta <https://consciousnessonline.wordpress.com/2012/02/17/dissolving-the-hard-problem-of-consciousness/>
- Chalmers, D. J. 2014. Intuitions in philosophy: a minimal defense. *Philosophical Studies*, 171(3), 535–544.
- Chalmers, D. J. 2018a. The meta-problem of consciousness. *Journal of Consciousness Studies*, 25(9–10), 6–61.
- Chalmers, D. J. 2018b. *The meta-problem of consciousness* [video]. *The Science of Consciousness Conference, Tucson 6.4.2018*. Noudettu 12.12.2018 osoitteesta <https://www.youtube.com/watch?v=E9xbQXqgAA>
- Churchland, P. M. 1989. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. London: MIT Press.
- Churchland, P. M. 1996. The rediscovery of light. *Journal of Philosophy*, 93(5), 211–228.
- Churchland, P. S. 1996. The hornswoggle problem. *Journal of Consciousness Studies*, 3(5–6), 402–408.
- Clark, A. 2000. A case where access implies qualia? *Analysis*, 60(1), 30–37.
- Clark, A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03), 181–204.
- Clark, A. 2015. Embodied prediction. Teoksessa T. Metzinger & J. M. Windt (toim.), *Open MIND* [online]. MIND Group, 263–283. Noudettu 15.12.2018 osoitteesta <http://cognet.mit.edu/pdfviewer/book/9780262340281/chap20>
- Clark, T. W. 1995. Function and phenomenology: Closing the explanatory gap. *Journal of Consciousness Studies*, 2(3), 241–254.
- Crick, F., & Koch, C. 2003. A framework for consciousness. *Nature Neuroscience*, (6), 119–126.
- Dennett, D. C. 1978. Why you can't make a computer that feels pain. *Synthese*, 38(3), 415–456.
- Dennett, D. C. 1979. On the absence of phenomenology. Teoksessa D. F. Gustafson & B. L. Tapscott (toim.), *Body, Mind, and Method*. Dordrecht, Holland: D. Reidel

- Publishing Company, 93–113.
- Dennett, D. C. 1984. Carving the mind at its joints. *Contemporary Psychology: A Journal of Reviews*, 29(4), 285–286.
- Dennett, D. C. 1988. Quining qualia. Teoksessa D. J. Chalmers (toim.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press, 226–246.
- Ilmestynyt ensimmäisen kerran teoksessa Anthony J. Marcel & E. Bisiach (toim.) *Consciousness in contemporary science*. Oxford University Press.
- Dennett, D. C. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.
- Dennett, D. C. 1995. The unimagined preposterousness of zombies. *Journal of Consciousness Studies*, 2(4), 322–326.
- Dennett, D. C. 1996. Facing backwards on the problem of consciousness. *Journal of Consciousness Studies*, 3(1), 4–6.
- Dennett, D. C. 1998. The myth of double transduction. Teoksessa S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (toim.), *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, 97–107.
- Dennett, D. C. 2001a. Are we explaining consciousness yet? *Cognition*, 79(1–2), 221–237.
- Dennett, D. C. 2001b. The fantasy of first-person science. Noudettu 3.1.2019 osoitteesta <https://ase.tufts.edu/cogstud/dennett/papers/chalmersdeb3dft.htm>
- Dennett, D. C. 2003. Who's on first? Heterophenomenology explained. *Journal of Consciousness Studies*, 10(9–10), 19–30.
- Dennett, D. C. 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge (MA): MIT Press.
- Dennett, D. C. 2006. Higher-order truths about chess. *Topoi*, 25(1–2), 39–41.
<https://doi.org/10.1007/s11245-006-0005-2>
- Dennett, D. C. 2009a. Darwin's strange inversion of reasoning. *Proceedings of the National Academy of Sciences*, 106(Suppl 1), 10061–10065.
<https://doi.org/10.1073/pnas.0904433106>
- Dennett, D. C. 2009b. Intentional Systems Theory. Teoksessa B. McLaughlin, A. Beckermann, & S. Walter (toim.), *Oxford Handbook of Philosophy of Mind*. Oxford University Press, 339–350.
- Dennett, D. C. 2012. The mystery of David Chalmers. *Journal of Consciousness Studies*, 19(1–2), 86–95.
- Dennett, D. C. 2013. *Intuition Pumps and Other Tools for Thinking*. Lontoo: Penguin

Books.

- Dennett, D. C. 2015a. How our belief in qualia evolved, and why we care so much - a reply to David H. Baßler. Teoksessa T. Metzinger & J. M. Windt (toim.), *Open MIND* [online]. MIND Group, 425–429. Noudettu 15.12.2018 osoitteesta <http://cognet.mit.edu/pdfviewer/book/9780262340281/chap31>
- Dennett, D. C. 2015b. Why and how does consciousness seem the way it seems? Teoksessa T. Metzinger & J. M. Windt (toim.), *Open MIND* [online]. MIND Group, 401–411. Noudettu 15.12.2018 osoitteesta <http://cognet.mit.edu/pdfviewer/book/9780262340281/chap29>
- Dennett, D. C. 2016. Illusionism as the obvious default theory of consciousness. *Journal of Consciousness Studies*, 23(11–12), 65–72.
- Dennett, D. C. 2017. *From Bacteria to Bach and back: The Evolution of Minds*. New York; London: W.W. Norton & Company.
- Dennett, D. C., & Steglich-Petersen, A. 2008. *The Philosophical Lexicon*. Noudettu 13.12.2018 osoitteesta <https://www.philosophicallexicon.com/>
- Farrell, J. 2016. ‘What it is like’ talk is not technical talk. *Journal of Consciousness Studies*, 23(9–10), 50–65.
- Frankish, K. 2007. The anti-zombie argument. *The Philosophical Quarterly*, 57(229), 650–666.
- Frankish, K. 2016a. Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23(11–12), 11–39.
- Frankish, K. 2016b. Not disillusioned: Reply to commentators. *Journal of Consciousness Studies*, 23(11–12), 256–289.
- Greenberg, W. J. 1998. On Chalmers’ ”principle of organizational invariance” and his ”dancing qualia” and ”fading qualia” thought experiments. *Journal of Consciousness Studies*, 5(1), 53–58.
- Haugeland, J. 1985. *Artificial intelligence: The very idea*. Cambridge (MA): The MIT Press.
- Havlík, M., Kozáková, E., & Horáček, J. 2017. Why and how. The Future of the central questions of consciousness. *Frontiers in Psychology*. Noudettu 30.1.2019 osoitteesta <http://journal.frontiersin.org/article/10.3389/fpsyg.2017.01797/full>
- Hintikka, J. 1999. The Emperor’s New Intuitions. *Journal of Philosophy*, 96(3), 127–147.
- Hobson, J. A., & Friston, K. J. 2014. Consciousness, dreams, and inference: The

- Cartesian Theatre revisited. *Journal of Consciousness Studies*, 21(1–2), 6–32.
- Irvine, E. 2017. Explaining what? *Topoi*, 36(1), 95–106.
- Jackson, F. 1982. Epiphenomenal qualia. *The Philosophical Quarterly*, 32, 127–136.
- Jackson, F. 1986. What Mary didn't know. *The Journal of Philosophy*, 83(5), 291–295.
- Jackson, F. 2003. Mind and illusion. *Royal Institute of Philosophy Supplement*, 53, 251–271. <https://doi.org/10.1017/S1358246100008365>
- Kammerer, F. 2018. Can you believe it? Illusionism and the illusion meta-problem. *Philosophical Psychology*, 31(1), 44–67.
- Kim, J. 2005. *Physicalism, or Something Near Enough*. Princeton University Press.
- Kirk, R. 1974. Sentience and behaviour. *Mind*, LXXXIII(329), 43–60.
- Kolers, P. A., von Grünau, M. 1976. Shape and color in apparent motion. *Vision Research*, 16(4), 329–335.
- Kripke, S. 1972/1980. *Naming and Necessity*. 1980 p. Oxford: Basil Blackwell.
Julkaistu ensimmäisen kerran teoksessa G. Harman & D. Davidson (toim.) *Semantics of Natural Language*. Dordrecht: Reidel, 253–355 (1972).
- Levine, J. 1983. Materialism and qualia: The explanatory gap. *Pacific philosophical quarterly*, 64(4), 354–361.
- Levine, J. 1994. Out of the closet: A qualophile confronts qualophobia. *Philosophical Topics*, 22(1–2), 107–126.
- Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford University Press.
- Levine, J. 2009. The explanatory gap. Teoksessa B. McLaughlin, A. Beckermann, & S. Walter (toim.), *Oxford Handbook of Philosophy of Mind*. Oxford University Press, 281–291.
- Lewis, D. 1995. Should a materialist believe in qualia? *Australasian Journal of Philosophy*, 73(1), 140–144.
- Loar, B. 1990. Phenomenal states. *Philosophical Perspectives*, 4(Action Theory and Philosophy of Mind), 81–108.
- Loorits, K. 2019. *Phenomenal consciousness as a structure in our brains*. University of Helsinki. Väitöskirja.
- MacLennan, B. J. 1996. The elements of consciousness and their neurodynamical correlates. *Journal of Consciousness Studies*, 3(5), 409–424.
- Madary, M. 2015. Extending the explanandum for predictive processing - A commentary on Andy Clark. Teoksessa T. Metzinger & J. M. Windt (toim.), *Open MIND* [online]. MIND Group, 285–295. Noudettu 15.12.2018 osoitteesta

<http://cognet.mit.edu/pdfviewer/book/9780262340281/chap21>

- Mandik, P. 2016. Meta-illusionism and qualia quietism. *Journal of Consciousness Studies*, 23(11–12), 140–148.
- Marinsek, N. L., & Gazzaniga, M. S. 2016. A split-brain perspective on illusionism. *Journal of Consciousness Studies*, 23(11–12), 149–159.
- Mashour, G. A., & LaRock, E. 2008. Inverse zombies, anesthesia awareness, and the hard problem of unconsciousness. *Consciousness and Cognition*, 17(4), 1163–1168.
- Matey, J. 2012. Comments on Schier and Carruthers: Dissolving the hard problem of consciousness. Noudettu 20.2.2019 osoitteesta <https://consciousnessonline.wordpress.com/2012/02/17/dissolving-the-hard-problem-of-consciousness/>
- Mørch, H. H. 2018. The evolutionary argument for phenomenal powers. *Philosophical Perspectives*, 31(1), 293–316.
- Nagel, T. 1970. Armstrong on the mind. *The Philosophical Review*, 79(3), 394–403.
- Nagel, T. 1974. What is it like to be a bat? *The Philosophical Review*, 83(4), 435–450.
- Papineau, D. 2002. *Thinking about consciousness*. Oxford University Press.
- Papineau, D. 2006. Phenomenal and perceptual concepts. Teoksessa T. Alter & S. Walter (toim.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, 111–144.
- Papineau, D. 2011. What exactly is the explanatory gap? *Philosophia*, 39(1), 5–19.
- Papineau, D. 2017. Competence without comprehension: David Papineau on the peculiar philosophical assumptions of Daniel Dennett. The Times Literary Supplement. Noudettu 22.1.2019 osoitteesta <https://www.the-tls.co.uk/articles/public/competence-without-comprehension-dennett/>
- Pauen, M. 2017. The Functional Mapping Hypothesis. *Topoi*, 36(1), 107–118.
- Pereboom, D. 2011. *Consciousness and the Prospects of Physicalism*. Oxford University Press.
- Pereboom, D. 2016. Illusionism and anti-functionalism about phenomenal consciousness. *Journal of Consciousness Studies*, 23(11–12), 172–185.
- Piccinini, G. 2017. Access denied to zombies. *Topoi*, 36(1), 81–93.
- Preston, A. 2015. Intuition in analytic philosophy. *Teorema*, XXXIV(3), 37–55.
- Price, H., & Corry, R. (toim.). 2007. *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford University Press.
- Price, M. C. 1996. Should we expect to feel as if we understand consciousness? *Journal*

- of Consciousness Studies*, 3(4), 303–312.
- Putnam, H. 1975. The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Raatikainen, P. 2018. Chalmersin argumentti materialismia vastaan. *Ajatus*, 75(1), 13–56.
- Rao, R. P. N., & Ballard, D. H. 1999. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87.
- Revonsuo, A. 2010. *Consciousness: The science of subjectivity*. Psychology Press.
- Rey, G. 2016. Taking consciousness seriously – as an illusion. *Journal of Consciousness Studies*, 23(11–12), 197–214.
- Robinson, H. 1993. Dennett on the knowledge argument. *Analysis*, 53(3), 174–177.
- Russell, B. 1910. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, 11(5), 108–128.
- Schneider, S. 2007. Daniel Dennett on the nature of consciousness. Teoksessa M. Velmans & S. Schneider (toim.), *The Blackwell Companion to Consciousness*. Blackwell Publishing, 313–324.
- Searle, J. R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(03), 417–457.
- Searle, J. R. 1992. *The Rediscovery of the Mind*. MIT Press.
- Searle, J. R. 2004. *Mind: A Brief Introduction*. Oxford University Press.
- Seth, A. K. 2015. The cybernetic bayesian brain - from interoceptive inference to sensorimotor contingencies. Teoksessa T. Metzinger & J. M. Windt (toim.), *Open MIND* [online]. MIND Group, 1451–1474. Noudettu 15.12.2018 osoitteesta <http://cognet.mit.edu/pdfviewer/book/9780262340281/chap104>
- Seth, A. K. 2016. The real problem. Noudettu 24.1.2019 osoitteesta <https://aeon.co/essays/the-hard-problem-of-consciousness-is-a-distraction-from-the-real-one>
- Smart, J. J. C. 1959. Sensations and brain processes. *The Philosophical Review*, 68(2), 141–156.
- Spurrett, D., & Cowley, S. 2010. The extended infant: Utterance activity and distributed cognition. Teoksessa R. Menary (toim.), *The Extended Mind*. MIT Press, 295–323.
- Strawson, G. 2017. A hundred years of consciousness: "a long training in absurdity". *Estudios de Filosofía*, 59(Enero-junio), 9–43. Esitelmä konferenssissa Isaiah Berlin Lectures, Wolfson College, Oxford, 25.5.2017.

- Sundström, P. 2008. Is the mystery an illusion? Papineau on the problem of consciousness. *Synthese*, 163(2), 133–143.
- Sytsma, J., & Machery, E. 2010. Two conceptions of subjective experience. *Philosophical Studies*, 151(2), 299–327.
<https://doi.org/10.1007/s11098-009-9439-x>
- Sytsma, J., & Machery, E. 2012. On the relevance of folk intuitions: A commentary on Talbot. *Consciousness and Cognition*, 21(2), 654–660.
<https://doi.org/10.1016/j.concog.2011.04.003>
- Tononi, G., & Koch, C. 2015. Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668). <https://doi.org/10.1098/rstb.2014.0167>
- Velmans, M. 2007. An epistemology for the study of consciousness. Teoksessa M. Velmans & S. Schneider (toim.), *The Blackwell Companion to Consciousness*. Blackwell Publishing, 711–725.
- Weinberg, J. M., Gonnerman, C., Buckner, C., & Alexander, J. 2010. Are philosophers expert intuiters? *Philosophical Psychology*, 23(3), 331–355.
- Wiese, W., & Metzinger, T. 2017. *Vanilla PP for philosophers: A primer on predictive processing*. MIND Group. <https://doi.org/10.15502/9783958573024>
- Yudkowsky, E. 2008. Zombies! Zombies? Noudettu 18. tammikuuta 2019, osoitteesta <https://www.lesswrong.com/posts/fdEWWr8St59bXLbQr/zombies-zombies>
- Zahavi, D. 2007. Killing the straw man: Dennett and phenomenology. *Phenomenology and the Cognitive Sciences*, 6(1–2), 21–43.