



Komputationalismi, eksternalismi ja Descartesin ”objektiivinen realiteetti”

JANI SINOKKI

Johdanto

Tavoitteeni on yhdistää kuvaus mielestä laskennallisena eli *komputaationaalisena* prosessina laajempaan viitekehykseen, joka selittää myös mielen *intentionaaliset* tai merkitysrelaationaaliset ominaisuudet luontevalla ja johdonmukaisella tavalla. *Komputationalismi* on se tunnettu tai pahamaineinen näkemys – riippuu, keneltä kysyy – joka olettaa mielen toiminnan olevan kuvattavissa laskentana ja periaatteessa kuin *tietokoneohjelmana*. *Intentionaalisuus* on taas ihmismielen usein perustavaksi katsottu suuntautuneisuus johonkin itsensä ulkopuoliseen, tai mielen kyky esittää maailman ja sen olioiden olevan tällä tai tuolla tavalla.

Komputationalismin mukaan mielen operaatioihin vaikuttavat ominaisuudet ovat kaikki luonteeltaan ”syntaktisia” tai muodollisia piirteitä, siis ominaisuuksia, jotka ovat relevantteja tiettyjen laskennallisten operaatioiden kannalta. Mikään, mikä ei ilmene näiden syntaktisten ominaisuuksien tasolla, ei komputationalismin näkökulmasta voi vaikuttaa mielen prosesseihin millään tavalla. Tämä aiheuttaa ongelman: miten merkitysrelaationaaliset ominaisuudet oikeastaan syntyvät päänsisäisten

laskennallisten operaatioiden tuloksena? Kuuluisien ajatuskokeiden mukaan (erit. Putnam 1975) kaksi ajattelijaa voi olla päänsisäisesti identtisisissä tiloissa, mutta ympäristössä olevien havainnon ulottumattomissa olevien erojen vuoksi (esim. ero aineiden kemiallisen rakenteen tasolla) silti ajatella eri ajatuksia.¹ Komputationalismin tapauksessa ongelma on se, että mitä ilmeisimmin kaksi ajattelijaa voi olla ajatusprosesseiltaan identtisisissä syntaktisissa tiloissa ja silti ajatella eri ajatuksia. Koska ainoa ilmeinen tapa, jolla komputationalismi voisi ottaa huomioon myös merkitysrelaationaaliset erot ajattelijoiden ympäristöissä, olisi se, että ne tulisivat ilmi jo mielensisäisten syntaktisten ominaisuuksien tasolla, näyttää niin kutsutun eksternalismin haaste erityisen suurelta juuri komputationalismille. Olettavasti nuo merkitysrelaatiot näet *eivät* ilmene osana oletettuja syntaktisia ominaisuuksia sen paremmin kuin viittaussuhteet kielen syntaktisten ominaisuuksien tasolla tai totuus osana väitelauseiden loogista muotoa.²

¹ Putnam itse tosin keskittyi kielellisiin merkityksiin, mutta muut laajensivat hänen alkuperäiset ajatuksensa koskemaan myös ”mielensisältöjä” (ks. esim. Burge 1979).

² Väitelauseiden tapauksessa on tietysti huomattava se poikkeus, että tautologisten väitelauseiden totuus ilmenee jo loogisen muodon tasolla.

Useat komputationalistisesti orientoituneet filosofit kyllä väittävät laskennallisten tilojen olevan itsessään myös semanttisia tiloja (esim. Peacocke 1994; 1999). Tämä käsitys näyttää kuitenkin ongelmalliselta (ks. Piccinini 2004; 2008). Esimerkiksi Peacocken näkemys semanttisista komputationalistisista tiloista näyttää etenevän siitä oletuksesta, että jos funktionalinen kuvaus kognitiivisesta toiminnasta (esim. näköhavainto esineestä) sisältää myös komputationalistisen kuvauksen mielen toiminnasta, on niiden komputationalististen kuvausten sisällettävä myös semanttisia, relationaalisia ominaisuuksia (esim. suhde näköhavainnon kohteeseen). Kuten Peacocken ja hänen esitystään vastustavan Frances Eganin näkemysten vaihdosta käy ilmi (Egan 1999; Peacocke 1999), Peacocken semanttinen käsitys mielen tiloista näyttää oletettavan lähtökohdissaan, että mikäli tietty tila tai toiminto,

Esitän, että Descartesin "objektiivista realiteettia" koskeva näkemys saattaa antaa välineitä tämän ongelman jäsentämiseen ja ratkaisemiseen. Descartes näyttää käsittäneen mielen kuvaamisen olevan mahdollista sisäisen prosessin tasolla, jota myös komputationalismi tarkastelee. Toisaalta hän näyttää ajatelleen, että tuo "syvempi" kuvaus mielen mekaniikasta on kuitenkin vain osa laajempaa kuvausta, joka selittää myös mielen merkitysrelaatiot maailmaan, ja on alisteinen tuolle ylemmän tason kuvaukselle. Tämä näkökulma muistuttaa monin tavoin huomattavasti uudempaa näkemystä, jonka Jerry Fodor on esittänyt. Tarkempi vertailu kuitenkin paljastaa merkittäviä eroja näiden kahden näkemyksen välillä. Näiden erojen ymmärtäminen on erityisen tärkeää yhtäältä komputationalismin mahdollisuuksien ja toisaalta sen sisäisten rajoitusten ymmärtämiselle. Esitän, että Descartesia mukailevan lähestymistavan avulla voimme löytää tavan yhdistää komputationalistisen kuvauksen mielen mekaniikasta niihin olennaisiin merkitysrelaatioihin, joihin Putnam ja muut niin kutsutut eksternalistit ovat kiinnittäneet huomiota.

Pohjustakseni Descartes-tulkintaani ja sen seurauksia komputationalismille tarkastelen ensin komputationalismia hyvin yleisellä tasolla, ja esittelen erään keskeisen eksternalismia vaikeavaan ongelman. Tämä ongelma koskee sitä, että eksternalismia tukevat ajatuskokeet näyttävät edellyttävän sopivaa epätarkkuutta mielensisällöiltä, jotta nuo ajatuskokeet voisivat toimia. Esitän, että havaitsemalla tämän edellytyksen voimme huomata, että Putnamin eksternalististen ajatuskokeiden todistusvoima on vähintäänkin kyseenalainen, ja että pohjimmiltaan niiden eristämä ilmiö on hyvin arkipäiväinen ajattelun luonteseen lähtökohtaisesti kuuluva piirre. Näiden tarkastelujen tarkoitus on auttaa meitä ymmärtämään paremmin, mistä jännite

jota kuvataan komputationalistisesti, on osa mielen tai toimijan funktiota, joka on itsessään relationaalinen (esim. näköhavainto esineestä), myös kuvaus tuollaisesta tilasta on näin ollen relationaalinen. Kuten myöhemmin osoitan, tämä ajatus on sekaantunut. (Kiitän toista arvioijaa hyödyllisestä kysymyksestä koskien semanttisen eksternalismin ongelmia.)

komputationalismin ja eksternalismin välillä juontuu, ja miten voimme lähteä purkamaan tätä ongelmavyöhytää.

Mielen mekanismi – laskenta ja representaatiot

Ajatus, että mieli on pohjimmiltaan laskutoimituksia suorittava, ei ole erityisen uusi. Jo Thomas Hobbes ehdotti näkemystä *Leviathanissaan* (1651). Vaikutusvaltaisim täsmällisempi laskennan idea on kuitenkin peräisin Alan Turingilta (1935), ja se siis on paljon myöhempää perua kuin Hobbesin filosofia. Lisäksi myös laskutoimituksen käsitteen soveltamisala oli Hobbesilla suoraviivaisempi kuin millaisena se tänä päivänä usein käsitetään. Hobbesin ajatus laskutoimituksesta perustui aritmetiikalle, ja *Leviathanissa* hän kirjoittaa vain järjenkäytöstä aritmeettisena prosessina (tosin on joitakin syitä olettaa hänen ajatelleen kaike ajattelun olleen luonteeltaan laskentaa³).

Käsitykset komputaation täsmällisestä luonteesta vaihtelevat, mutta kirjoitukseni tarpeisiin riittää se yleinen luonnehdinta, että laskennallisten sääntöjen soveltaminen nähdään algoritmisena, eli yksiselitteisistä askeleista muodostuvana. Askeleet sinänsä voivat perustua myös todennäköisyyksille. Mikään ei siis pakota meitä käsittämään mieltä aritmeettisena laskukoneena, kuten Hobbes ehdotti, vaikka tämä kieltämättä helpottaisikin mielen mekanismin kuvaamista.

Keskityn tässä kirjoituksessa sellaiseen hyvin yleiseen näkemykseen mielestä, joka yhdistää monia komputationalistisia mielen malleja. Tämän näkemuksen mukaan mieli koostuu sisällöistä, *representaatioista*. Näillä mielensisällöillä on ainakin tiettyjä syntaktisia, muodollisia ominaisuuksia, joiden perusteella mieli käsittelee niitä tiettyjen (rekursiivisten) sääntöjen avulla. Kutsun niitä laskennallisia operaatioita, joilla mieli manipuloi representaatioita, lyhyesti *komputaatioiksi*.

³ Hobbes käsittelee mielen toimintoja laskentana myös *De Corpore*ssa (1655) ja soveltaa aritmeettista mallia siellä laajemmin kuin *Leviathanissa*.

Mielen komputaatiot ovat yleisesti ottaen abstrakti kuvaus ajattelun *toiminnallisesta mekanismista*. Eräs filosofeja jakava kysymys, joka liittyy tähän abstraktiin kuvaukseen, on abstraktin kuvauksen suhde siihen materiaan, joka *implementoi* tuon kuvauksen, kuten aivoihin tai – keinotekoisena ajattelun tapauksessa – tietokoneen loogiseen piiriin. Laskennallisen abstraktin mielen mallin ei tarvitse toteutua eli realisoitua sellaisenaan esimerkiksi neuronien toiminnassa tai loogisissa porteissa, *koska* sen on tarkoitus olla ylemmän tason kuvaus *mielen* toiminnasta. Voimme siis hyvin ajatella, että prosessit, joissa tarkastelumme kohteena oleva ylemmän tason mielen malli toteutuu fyysisissä aivoissa tai logiikkapiirissä, edellyttävät omaa (komputationaalista) kuvaustaan, ja että kysymys tämän implementaatiotason kuvauksen oikeellisuudesta on lopulta erillään kysymyksestä, joka koskee ylemmän tason mielen mallin kuvauksen oikeellisuutta. Tämä ajatus erosta mielen mallin ja sen partikulaaristen implementaatioiden välillä liittyy kiinteästi näkemykseen siitä, että ylemmän tason abstrakti malli mielestä käsitetään tyypillisesti *monitoteutuvaksi* (engl. *multiply realizable*), eli erilliseksi niistä täsmällisistä konfiguraatioista, jotka voivat olla tuon mallin fyysinen implementaatio. En tarkastele kysymystä mielen mallien implementaatioista sinänsä lainkaan, ja monitoteutuvuuden oletan todeksi argumentoimatta sen puolesta.

Mielen operaatiot aritmetiikasta järkeilyyn, uskomusten muodostamisesta todennäköisyyksien arviointiin ja erilaisten kokemuksen kokemiseen, voidaan komputationalistisen mallin mukaan käsittää mielensisällöille eli representaatioille tehtävinä komputaatioina. Olipa käsitys representaatioista mikä tahansa, näyttää perustellulta olettaa, että *mikäli* mielen operaatiot ovat jonkinlaisia laskutoimituksia, *niin* mielensisällöillä täytyy olla tiettyjä muodollisia, formaaleja piirteitä, joita voidaan luonnehtia representaatioiden *syntaktisiksi ominaisuuksiksi*. Näin on siksi, että laskennan voidaan ajatella muodostuvan loogisten sääntöjen soveltamisesta mielensisältöjen jäsentämiseen, ja loogiset säännöt tyypillisesti käsitetään luonteeltaan muodollisiksi, puhtaasti syntaktisiksi yhdistelysäännöiksi. Toisaalta

olisi hyvin vaikea nähdä mitä ei-syntaktisille ominaisuuksille perustuva komputaatio voisi edes olla.

Tässä lienee vahvin syy myös kannattaa komputationalismia jossain muodossa. On hyvin vaikea kuvitella, miten mikään eilaskennallinen malli mielen toiminnasta voisi kuvata ja selittää ajattelun *systemaattisia* kognitiivisia ilmiöitä, kuten kielen oppimista ja käyttämistä, loogista päättelykykyä, mielen periaatteessa rajatonta mahdollisuutta tuottaa ja käsittää uusia propositionia tai mielen kykyä soveltaa vanhaa tietoa täysin uusiin ilmiöihin. Mikään heuristinen tai ”laadullinen” malli ei voisi selittää näiden ilmiöiden systemaattisuutta eikä tämän systemaattisuuden johdonmukaista ilmentymistä ihmisajattelussa.

Mikäli komputaatio käsitetään syntaktisten sääntöjen soveltamiseksi, näyttää siltä, että representaatioiden komputaation kannalta olennaiset piirteet ovat niitä syntaktisia ominaisuuksia, joiden perusteella mieli soveltaa syntaktisia sääntöjä representaatioihin. On kuitenkin selvää, että pelkkä syntaksi ilman semantiikkaa olisi riittämätön kuvaus ajattelusta. Mieli ei pelkästään manipuloi tyhjiä skemaattisia symboleja ja kaavoja, vaan on toimintansa vuoksi suhteessa maailmaan ja sen olioihin. Niinpä näkemys mielen toiminnasta komputaatioina edellyttää myös jotakin selitystä siitä, miten mieli kykenee viittaamaan itsensä ulkopuolelle, tai vaihtoehtoisesti, miten representaatiot, joilla on tiettyjä syntaktisia ominaisuuksia, ovat lisäksi informaatiota maailmasta ja sen olioista.

Esimerkiksi Hobbessin näkemyksessä olennaista on, että ajattelu koostuu sisäisistä ideoista, joita voidaan ilmaista kielellisesti. Kielen ensisijainen merkitys on toimia merkkeinä, joiden avulla ajattelija pystyy tuomaan mieleen tiettyjä ideoita tai käsitteitä. Vasta kielen toissijainen merkitys on puhujien välisessä kommunikaatiossa – sanojen avulla puhujat voivat koordinoita mielissään olevia ideoita keskenään. Hobbes esittääkin aritmeettisen kuvauksen mielestä nimenomaan sanojen yhdistelysääntöinä. Hobbessin näkemyksestä kielellisen viittaamisen eli *signifikaation* luonteesta on tosin erimielisyyttä, mutta esimerkiksi John Locke myöhemmin toteaa eksplisiittisesti, että sanat viittaavat vain ja ainoastaan mielen representaatioihin eli

ideoihin, ja sanojen viittaussuhde maailman olioihin – siis se, mitä nykyään kutsutaan kielifilosofiassa *referenssiksi* – on johdannaista ideoiden sellaisista ominaisuuksista, jotka ovat olennaisia mielen ulkoisten olioiden esittämiselle (Locke 1689/1975). Tällaisessa näkemyksessä representaatioiden (mahdollisen) viittaussuhteen mielen ulkopuolisiin olioihin ajatellaan muodostuvan kuvauksellisesta eli deskriptiivisestä mielensisällöstä yhdistettynä tämän kuvauksen tyydyttymisehtoihin. Representaatio on tällöin aina sen olion representaatio, jota representaatio tarkimmin vastaa mielen ulkoisessa todellisuudessa, ja mikä olio siis ”tyydyttää” representaation sisältämän kuvauksen.⁴

Oleennaista tässä *kuvaus plus tyydyttymisehdot* -näköyksessä, jota hyvin todennäköisesti ainakin Locke ja mahdollisesti myös Hobbes kannatti mielensisältöjen semanttisista ominaisuuksista, on, että representaatioiden informaatioisisältö on puhtaasti mielensisäistä. Ainoastaan informaation tyydyttymisehdot (ja mahdolliset tarkkuus- tai totuusehdot) ovat relationaalisia ja edellyttävät tiettyä suhdetta maailmaan. Kuitenkin representaatioita sinänsä voidaan, ja ehkä pitääkin, tarkastella vain mielen sisäisinä olioina, eikä niistä teoretisointi sinänsä edellytä minkään mielenulkoisen huomioimista. Kun tämä yhdistetään Hobbesin kannattamaan ajatukseen komputaatiosta mielen mekanismina, seuraa näkemys, jonka mukaan sekä mielen komputationaalinen mekanismi ja representaatioiden syntaktiset ominaisuudet että myös representaatioiden semanttiset tai informationaaliset ominaisuudet ovat kaikki tarkasteltavissa puhtaasti mielen sisäisinä seikkoina. Voidaan siis sanoa, että ajattelun luonteen ymmärtäminen ei tämän näkemyksen puitteissa edellytä minkään mielen ulkopuolisen huomioon ottamista, mahdollisesti pois lukien totuus- tai tyydyttymisehdot,

⁴ Esimerkiksi Arnauld lukee Descartesin idea-teoriaa tällä tavalla, mutta Descartes kieltää tämän tulkinnan oikeellisuuden (ks. *Neljännet vastaväitteet* ja Descartesin vastaukset niihin; Descartes 2002, 162–202; CSM2, 138–178; AT7, 196–256).

jotka joka tapauksessa ovat relationaalisia luonteeltaan. Ajattelu on siis sekä mekanismeiltaan että myös merkityksiltään *ensisijaisesti mielen sisäinen prosessi*.

Eksternalismi

Edellä tarkasteltu *internalistinen* näkemys ajattelusta mielen sisäisenä prosessina on näyttänyt monista puutteelliselta. Jo John Stuart Mill (*System of Logic*, I.v.1.,) erotti mielen sisäisen prosessin niistä ulkoisista suhteista, joissa mieli on ajattelun kohteisiin. Hän piti sisäisen prosessoinnin ja mielen relationaalisten ominaisuuksien samastamista erheenä. Viitaten mahdollisesti muiden muassa Hobbesiin ja Lockeen Mill esitti ”muutoin merkittävien” ajattelijoiden päätyneen läpeensä virheellisiin käsityksiin mielestä, koska he eivät riittävästi erottaneet kysymyksiä mielen (tai aivojen) sisäisen prosessin yksityiskohdista kysymyksistä, jotka koskivat mielen relationaalisuutta. Ajattelu on hänen mukaansa ensisijaisesti tiedollisessa relaatiossa olemista johonkin mielenulkoiseen. Esimerkiksi uskomukseni, että jää on vettä, on relaatio minun sisäisen uskomuksellisen tilani ja maailmassa vallitsevan asiointilan välillä.

Mill siis ajatteli, että ainakin propositionaalinen mielensisältö oli välttämättä kuvattava relationaalisesti. Nykykeskustelussa tätä propositionaalista relationaalista mielensisältöä kutsutaan *intentionaaliseksi* mielensisällöksi, ja kun erotan representaatioiden syntaktiset ominaisuudet niiden semanttisista ominaisuuksista, tarkoitan semanttisilla ominaisuuksilla juuri intentionaalista sisältöä. Pelkästään semanttisen tai intentionaalisen sisällön oletaminen sinänsä ei ota kantaa siihen, onko tuo sisältö välttämättä relationaalista. *Eksternalismina* tunnettu näkemys väittää näin olevan: sen mukaan ainakin osa mielensisältöjen semanttisista ominaisuuksista edellyttää relationaalista kuvausta. Toisin sanoen eksternalismin mukaan pelkkä sisäinen internalistinen, ei-relationaalinen kuvaus mielen sisällöstä ei riitä mielen täydelliseen kuvaamiseen.

Hilary Putnamin kuuluisat kaksoismaa-ajatuskokeet haastavat näkemykset, joiden mukaan ajattelu ei edellytä relationaalisten seikkojen huomioimista (Putnam 1975). Hän pyytää meitä muun muassa kuvittelemaan Oskarin ja kaksois-Oskarin, jotka ovat identtiset toistensa "vastineet" kahdessa eri maailmassa, meidän Maassamme ja Kaksoismaassa. Oskarit ovat molekyyli molekyyliä fyysisesti identtisiä, sisäisiltä mielentiloiltaan psykologisesti identtisiä ja he puhuvat samaa kieltä käyttäen sitä identtisillä tavoilla. Ainut ero, jonka saamme näiden kahden täsmälleen samanlaisissa ympäristöissä ja sisäisissä tiloissa oleville Oskareille kuvitella, on seuraava: se neste, johon molemmat viittaavat vetenä, ja joka on kummankin ympäristössä hanoista tuleva janon sammuttava neste, jota sataa taivaalta ja jota virtaa joissa ja merissä, on Maassa kemialliselta rakenteeltaan H₂O, kun taas kaksoismaassa se on identtiseltä näyttävää ja kemiallisesti identtisesti käyttäytyvää, mutta kemialliselta rakenteeltaan tyystin erilaista ainetta, XYZ:aa. Kun Oskarit katsovat edessään olevaa vedellä tai "kaksoisvedellä" täytettyä juomalasia, heidän ajatuksensa on muotoa "tuossa on vettä". Heidän sisäisissä psykologisissa tiloissaan ei siis ole mitään eroa. Sisäisestä identtisyydestä huolimatta meidän Oskarimme ajattelema ajatus on H₂O-ajatus ja kaksois-Oskarin ajattelema ajatus on XYZ-ajatus – he ajattelevat eri asioita, ja siis eri ajatuksia!

Ajatuskokeen tarkoitus on osoittaa, että pelkkä sisäinen psykologinen tila ei yksinään määrää, mikä ajatus on kyseessä. Putnamin käyttämät scifiesimerkit tuottavat lisämysterin tuntua ajatuskokeiden lopputulokseen, mutta ajatuskokeiden periaate on yllättävän arkinen. Esimerkiksi jääveistosnäyttelyssä voisin tietämättäni katsella lasista valmistettua huijausta, ja edessäni olevaa veistosta koskeva sisäinen mielentilani olisi täsmälleen sama kuin se olisi ollut, vaikka kyseessä olisi ollut oikea jääveistos. Silti lasiveistosta katsellessani minä ajattelen "lasiveistosajatusta" ja jääveistosta katsellessani "jääveistosajatusta". Ilmiö ei ole sinänsä niin erikoinen kuin Putnamin ajatuskokeita koskevan myöhemmän keskustelun perusteella voisi kuvitella,

ja se on suoraa seurausta siitä, että ajattelu on tyypillisesti jonkin kohdistuvaa eli intentionaalista. Tullessani huijatuksi jääveistosnäyttelyssä uskomukseni epätotuus johtuu ajattelemani veistoksen todellisesta rakenteesta. Jos uskon veistoksen olevan jäätä, mutta se onkin lasia, on minulla virheellinen käsitys edessäni olevasta esineestä. Tämä ilmiö näyttää puolestaan olevan hyvin arkinen ajattelun piirre: ajatuksen, kuten lauseiden tai propositioiden, totuus on relationaalista niiden tarkoitamiin kohteisiin. On triviaalia todeta, että ajatusten sen paremmin kuin lauseidenkaan "tosiasioidenmukaisuus" ei ole niiden sisäinen ominaisuus, vaan riippuu siitä, miten tarkoitetut asiat todellisuudessa ovat.

Eksternalististen ajatuskokeiden onnistuminen näyttääkin edellyttävän, että käytetyt uskomukset ovat riittävässä määrin epätarkkoja ja yleisluontoisia. Putnam näyttää tunnistavan tämän, koska hän täsmällisesti edellyttää kaksoismaa-ajatuskokeessa, että Oskarien kuvitellaan elävän *ennen* nykyisen kemian kehittymistä.⁵ Koska ympäristöjen välillä on merkittävä kemiallinen ero, eivät Oskarit voisi olla identtisissä sisäisissä tiloissa enää vedeksi kutsutun aineen kemiallisen rakenteen löytymisen jälkeen. Tämä rajoite uskomusten tarkkuuteen aiheuttaa myös sen, että Oskarien uskomukset ovat hyvin epätarkkoja. Jos oletamme, että meidän Oskarimme käsitys vedestä on tätä tarkempi ja sisältää esimerkiksi oletuksia veden kemiallisesta rakenteesta, muuttuu ajatuskokeen vangitsema ilmiö täysin arkiseksi. Jos meidän Oskarimme joutuisi teleportatuksi kaksoismaahan (scifiesimerkeissä mikä vain on mahdollista), hän ajattelisi juomalasin nähdessään "tuossa on vettä", mutta olisi yksinkertaisesti vain väärässä, koska hänen edessään ei ole H₂O:ta, vaan XYZ:ta. Hän olisi tuolloin vastaavassa tilanteessa kuin

⁵ Kuten aiemmin totesin, Putnamin oma huomio kohdistuu pikemmin kieleen kuin ajatteluun, joten tässä esitetty ongelma ei välttämättä koske hänen tosiasiallista näkemystään aivan samalla tavalla kuin se koskee niitä näkemyksiä, jotka koskevat "mielensisältöjen" eksternalistisuutta (esim. Peacocke 1994).

missä minä olen jääveistosnäyttelyssä tullessani huijatuksi lasiveistoksella. Näyttääkin itseasiassa siltä, että eksternalismin tukeksi esitettyjä ajatuskokeita ei voida toisintaa millään sellaisilla uskomuksilla tai ajatuksilla, jotka ovat tarkempia ja sisältörikkaampia suhteessa kohteisiinsa.

On myös huomattava, että on vaikea sanoa kuinka uskottavia ajatuskokeet todella ovat – niiden uskottavuus on haastettu lukuisia kertoja erilaisilla perusteilla (esim. Boghossian 1997; Unger 1984; Loar 1988; Fodor 1987; Horowitz 2001), joita ei ole tarpeen kerrata tässä. Keskeistä on huomata, että uskomuksia tai ajatuksia (tai niin kutsuttuja ”mielensisältöjä”) on ylipäänsä vaikea käsittää riippumattomiksi siitä kontekstista, jossa nuo uskomukset tai ajatukset aktualisoituvat. Tästä ei kuitenkaan seuraa, etteikö puhtaasti sisäistä tilaa voitaisi mielekkäällä tavalla erottaa kontekstiin kuuluvista asioista, kuten ajattelun kohteista, päinvastoin kuin Putnam näyttää ajattelevan.

Esimerkiksi Putnamin ajatuskokeen vesi voitaisiin korvata tiikereillä, jolloin Oskarit olisivat tiikeri-ajatuksia ajatellessaan identtisissä sisäisissä tiloissa, mutta selvästi ajattelisivat eri asioita: meidän Oskarimme ajattelisi Maa-planeetan evoluution tuloksena syntyneitä eläinlajeja (tai sen jäseniä), ja kaksois-Oskari ajattelisi kaksoismaan evoluution synnyttämää lajia. Sen tavanomaisen oletuksen puitteissa, että evolutiivinen alkuperä on olennaista lajiin kuulumisen kannalta, ovat ajattelun kohteet erilliset, koska lajit ovat erilliset (meidän tiikerimme ja kaksoistiikerit ovat syntyneet eri paikoissa erillisten, mutta samanlaisen, evolutiivisten prosessien seurauksena).

Tämä vasta-ajatuskoe mahdollistaa kahden huomion tekemisen. Ensinnäkin ajatuksen sisältö näyttää riippuvan laajemminkin kontekstista kuin vain ajattelun kohteen ominaisuuksista. Vaikka kaikki paikalliset detaljit (ajattelijan sisäinen tila, ajatuksen kohteen ominaisuudet) pidettäisiin vakioina, eri kontekstit riittävät tuottamaan eri ajatuksia. Esimerkiksi Jerry Fodor on esittänyt tämän kaltaisen tulkinnan eksternalistien ajatuskokeista. Hänen mukaansa ajatusten merkitysrelaationaaliset ominaisuudet ovat pohjimmiltaan funktio sisäisestä tilasta kon-

tekstiin (Fodor 1987). Eri konteksti tuottaa näin ollen välttämättä eri ajatuksia, mutta tästä ei seuraa, että mielensisäinen tila ei tyhjentävästi määräisi ajatuksen semanttisia ja relationaalisia ominaisuuksia *kussakin kontekstissa*.

Toisekseen on epäselvää, voivatko ulkoisetkaan seikat Putnamin esittämällä tavalla määrätä sitä, mitä ajattelemme. Voidaan nimittäin kuvitella myös tapauksia, joissa ulkoiset erot ympäristössä eivät näytä vaikuttavan ajatuksen sisältöön. Ajatustemme merkitykset ovat riippuvaisia myös monista (päämme sisäisistä) taustaoletuksista, jotka vaikuttavat merkittävästi siihen, millainen *käsityksemme*⁶ varsinaisesti on. Esimerkiksi luonnonlajeja koskevat presuppositiot, joihin olemme sitoutuneet tietoisesti tai tiedostamattomasti, saattavat olla hyvin merkityksellisiä sen suhteen, mitä tarkalleen ottaen ajattelemme (olen käsitellyt näitä ”metafyysisiä sitoumuksia”

⁶ On tärkeää erottaa *käsitykset* ja *käsitteet*. Helmi saattaa uskoa, että tiikerit ovat raidallisia, ja Urho uskoa, että tiikerit eivät ole raidallisia. Helmi ja Urho jakavat tällöin tiikerin käsitteen, mutta heillä on eri käsitykset tiikereistä. Käsitteet ovat välttämättä jaettuina eli *intersubjektivisia*, käsitykset puolestaan ovat enemmän alttiita vaihtelulle, joka riippuu *subjektiivisista* eli henkilökohtaisista seikoista. Käsitteillä on myös selvä normatiivinen ulottuvuus: käsitteen omaava henkilö voi käyttää tai soveltaa käsitettä väärin. Käyttäessään käsitettä väärin henkilöllä on väärä käsitys käytetystä käsitteestä. Koska Urhon uskomus, että tiikerit eivät ole raidallisia, on epätosi uskomus tiikerien lajista, hänen käsityksensä tiikereistä on väärä, vaikka hänellä onkin sama tiikerin käsite kuin Helmillä.

Myös sillä, onko käsitys *tiedostettu* (esim. uskomus) vai *tiedostamaton* (esim. oletus, ennakkoluulo), on merkitystä: monet käsityksistä, jotka ohjaavat toimintaamme, ovat enemmän tai vähemmän tiedostamattomia. Tiedostettuja käsityksiä eli käsityksiä, joiden merkitysisällöstä olemme tietoisia, voidaan nimittää *näkemyksiksi*. Näin ollen henkilöllä saattaa olla käsitys tiikereistä, vaikka hänellä ei olisikaan näkemystä esimerkiksi siitä, ovatko tiikeri raidallisia vai eivät. Siis riippuen käsityksestä henkilö käyttää tiikerin käsitettä (enemmän tai vähemmän) oikein tai väärin, oli hänellä näkemystä tiikereistä tai ei.

toisaalla, ks. Sinokki 2017b). Jos Oskarit uskoisivat tai tiedostamattaan olisivat sitoutuneet siihen, että geneettinen perimä tai evolutiivinen historia *ei* ole olennaista tiikerien lajiin kuulumiselle, he ajattelisivat eri maailmoissa eri asioita – meidän Oskarimme ajattelisi meidän tiikereitämme, kaksois-Oskari ajattelisi kaksois-tiikereitä, jotka siis ovat evolutiivisesti eri lajit. Kuitenkaan heidän ajatustensa totuusarvot eivät olisi riippuvaisia siitä, kummassa paikassa he ovat. Jos heidän tiikeri-käsityksensä eivät sisällä mitään sellaista rajausta lajiin kuulumiselle, joka riippuisi ajateltujen eläinten historiallisesta sukulinjasta ja noiden lajien kehitysympäristöstä, molemmat onnistuisivat ajattelemaan tosia ajatuksia lähiympäristönsä raidallisesta eläinlajista riippumatta siitä, onko heidän ympäristössään tiikereitä vai kaksois-tiikereitä. Ajatuksen sisäisen epämääräisyyden vuoksi Oskarit onnistuvat ajattelemaan tosia ajatuksia yhdellä ja samalla sisäisellä mielentilalla riippumatta siitä, että ero ympäristöissä on kuitenkin merkittävä. Näyttää siis siltä, että ympäristöjen välisten erojen merkitys ajattelun semanttisille ominaisuuksille on kaikkea muuta kuin selvä.

Olemme nyt huomanneet, että yhtäältä eksternalismin väite päänsisäisen tilan riittämättömyydestä ajattelun semanttisten ominaisuuksien määrittämisessä perustuu sille edellytykselle, että päänsisäinen tila on itsessään epätarkka. Toisaalta on myös epäselvää, kuinka paljon ulkoiset asiointilat todella vaikuttavat ajattelun semantiikkaan kaikissa tilanteissa. Näin ollen voimme todeta, että kaksoismaa-ajatuskokeet eivät näytä riittävän osoittamaan, että päänsisäinen kuvaus olisi riittämätön kuvaus ajattelusta. On kuitenkin syytä huomata, että ajatuskokeet kyllä onnistuvat osoittamaan sen, että ajatus mielen intentionaalisten suhteiden, kuten sanojen merkitysten tai uskomusten totuuden määräytymisestä pelkkien sisäisten representaatioiden tyydyttymisehtojen perusteella on kyseenalainen. Ajattelun merkitysrelationaaliset ominaisuudet ovat selvästikin monimutkaisempi asia kuin ”kuvaus plus tyydyttymisehdot” -näkemys esittää. On siis syytä ajatella, että vastaavuus mielen sisäisen prosessin ja maailman olioiden välillä ei ole yksinkertainen tai

yksisuuntainen relaatio, jossa jompikumpi, pään sisäinen tila tai ulkoinen asiointi, määrää ajattelun merkitysrelaatiot.

Eksternalismin haaste komputationalismille

Voidaan ajatella, että Millin ja Putnamin korostama ero sisäisen prosessin ja ulkoisten relaatioiden välillä on samankaltainen kuin edellä toteamani ero mielen yleisen mallin ja sen aivotason implementaation välillä. Mielen sisäinen prosessointi on välttämätöntä ajattelulle, mutta se on lopulta vain kuvaus yhdestä mahdollisesta yleisemmän kognition mallin implementaatiosta. Yleinen kognition malli koskee tiedollisia suhteita, joissa ajatteleva subjekti on maailmaan. Nämä tiedolliset suhteet eivät välttämättä ole suoraviivaisesti kuvattavissa mielen sisäisinä operaatioina, aivan kuten komputationaalisen mallin syntaktiset operaatiot eivät välttämättä ole suoraviivaisesti aivojen tai logiikkapiirin suorittamia operaatioita. Malli kognitiivisista suhteista, samoin kuin malli mielen laskennallisista operaatioista, on erotettava noiden mallien mahdollisista implementaatioista.

Näyttää selvältä, että komputationaalisen mielen mallin puitteissa ainoastaan ne representaatioiden ominaisuudet, jotka ovat syntaktisia, ovat relevantteja mielen komputaatioiden kannalta. Mikään sellainen sisällöllinen ero, joka ei ilmene myös erona syntaktisen käytöksen tasolla, ei yksinkertaisesti vaikuta komputaatioiden suorittamiseen eikä siis ajatteluun mitenkään. Näin ollen kaikkien semanttisten erojen mielensisällöissä on ilmennyttävä myös syntaktisina eroina. Tämä on merkittävä ongelma komputationalismille yleisesti mutta erityisesti eksternalismiin yhdistettynä.

Ongelma on seuraava. Eksternalismi esittää, että mielen sisäisten representaatioiden semanttiset ominaisuudet ovat ainakin toisinaan relationaalisia, eli edellyttävät suhdetta ulkoisiin asiointiloihin. Kuitenkin näyttää ilmeiseltä, että erot ulkoisissa asiointiloissa voivat vaikuttaa kausaalisesti käytökseemme vain aiheuttamalla eroja päidemme sisäisissä mielen- ja aivotiloissa. Ja edellä todetun mukaisesti ainoat ajattelun kannalta relevantit ja siis myös käytökseemme kannalta relevantit sisäiset tilat ovat

syntaktisesti kuvattavia. Niinpä jotta niillä olisi mitään kausaalista vaikutusta ajatteluamme ja käytökseemme, myös eksternalismin olettamien relationaalisten sisältöjen näyttää olevan pakko palautua lopulta ei-relationaalisiin, syntaktisesti kuvattavissa oleviin mielensisältöihin. Siis jos hyväksymme komputationalismin, emme voi hyväksyä eksternalismia, ja päinvastoin. Miten siis voimme pitää kiinni siitä intuitiivisesta näkemyksestä, että ajattelu voidaan kuvata tarkasti ja systemaattisesti mieleemme prosessointina, ja siitä yhtä intuitiivisesta näkemyksestä, että relaatiot ulkoisiin oloihin ja asiointeihin ovat olennainen osa tuota prosessointia?

Tähän ongelmaan on törmätty joitakin kertoja kirjallisuudessa.⁷ Esimerkiksi Jerry Fodor, eräs merkittävimmistä internalismin ja komputationalismin puolustajista, tuli ajattelunsa kehittyessä siihen tulokseen, että tietyt mielen semanttiset ominaisuudet on välttämättä kuvattava eksternalistisesti – pelkkä mielensisäisen prosessin analysointi ei vangitse mielentoiminnalle olennaisia merkitysrelaatioita. Ilman noiden relaatioiden huomioon ottamista komputationalistinen malli ei ole täydellinen kuvaus mielestä. Mutta edellisen ongelman vuoksi ei myöskään ole ilmeistä, miten tämä olisi mahdollista tehdä.

Fodorin vastaus ongelmaan on näkemyksenä samankaltainen kuin se, jonka luin aiemmin Millin näkemykseksi: kognitiiviset relaatiot, joita subjektilla on ulkoisiin asiointeihin, ja joihin eksternalismi kiinnittää erityistä huomiota, ovat osa yleisempää, ylemmän tason kognition mallia. Komputationalistinen kuvaus mielestä puolestaan on osa tuon laajemman, ylemmän tason mallin implementaatiota subjektissa. Kuvaukset ovat siis laajuudeltaan erilaisia, ja alaltaan suppeamman komputationalistisen kuvauksen suhde laajempaan kognition malliin *ei ole* osa-kokonaisuus-suhde, vaan pikemmin implementaatio- tai *toteuttamissuhde*.

Vastaava suhde voisi tulla esiin kahden sellaisen kuvauksen välillä, jotka koskevat vaihtorahan palauttamista asiakkaalle

⁷ Kuten yllä alaviitteessä 2 totesin, ongelman tiedostaminen filosofien keskuudessa vaikuttaa puutteelliselta.

kaupassa. Taloudellisen kuvauksen puitteissa tapahtuma on relationaalinen: hyödykettä vastaan on tarjottu vaihdon välinettä (eli rahaa), ja koska käytetty vaihdon väline -järjestely ei mahdollista aina täsmällistä pyydetyn korvauksen antamista myyjäosapuolelle, täytyy tämän hyvittää erotus hinnassa ja annettussa rahasummassa palauttamalla sopivan suuruinen summa. Toisaalta sama toiminta, joka kaupallisesta ja sosiaalisesta näkökulmasta on vaihtorahan palauttamista, voitaisiin kuvata myös pelkkänä fysiologisena toimintana, jossa myyjäosapuoli tekee tiettyjä liikkeitä siirrellen vaihdon välineinä toimivia fyysisiä kappaleita paikasta toiseen. Mikään tässä fysiologisessa kuvauksessa ei ilman taustatietoja (esimerkiksi tietoa, että liikuteltavat kappaleet ovat nimenomaan vaihdon välineitä) paljasta, että kyse on nimenomaan vaihtorahan palauttamiseksi sosiaalisesti luokiteltavasta toiminnasta. Vastaavasti moni erilainen tilanne voi realisoida ”vaihtorahan palauttamisen”. Vaihdon välineen muodolla tai liikkeillä, joita kassahenkilö tekee, ei varsinaisesti ole mitään väliä sen kannalta, onko kyseessä vaihtorahan palauttaminen vai ei: vaihtorahan palautus on monitoteutuva, eli monet fyysisen kuvauksen tasolla toisistaan eroavat ilmiöt voivat olla vaihtorahan palauttamista. Edellisen kappaleen terminologiaan yhdistettynä: tarkasteltu fyysinen kuvaus myyjän ruumiinliikkeistä on vaihtorahan palautuksen yksi mahdollinen fyysinen implementaatio.

Putnamin näkemyksen mukaan myös relationaaliset semanttiset ominaisuudet ovat relevantteja mielen toiminnalle ja *a fortiori* myös komputationaaliselle kuvaukselle. Voimme nyt argumentin nimissä rinnastaa nuo semanttiset, relationaaliset ominaisuudet talousteoreettisiin ominaisuuksiin ja komputationalistisen kuvauksen mielestä myyjän toiminnan fysikaaliseen kuvaukseen. Putnamilainen ehdotus vastaa ehdotusta, että koska taloudelliset ominaisuudet ovat olennainen osa tapahtuman luonnetta, eikä myyjän toimintaa voida ymmärtää asettamatta sitä laajempaan taloudellis-sosiaaliseen kontekstiin, on tuohon fysikaaliseen kuvaukseen sisällyttävä myös taloudellis-sosiaalisia ominaisuuksia. Tämä ehdotus puolestaan vai-

kuttaa järjettömältä ja näyttää perustuvan ilmiselvälle sekaan-
nukselle: kuvauksen tasot ovat sikäli erillisiä toisistaan, että ta-
loudellis-sosiaalinen kuvaus edellyttää vain *jonkin* fyysikaalisen
toteutumisen, mutta ei määrää (determinoi) mitään yhtä tiettyä
fyysikaalisen toteutumisen muotoa (esimerkiksi taloudellisesti
identtiset tilanteet voisivat toteutua monilla eri tavoin riippuen
kontekstista ja käytetystä maksuvälineestä – oravannahkojen
käsitteily edellyttäneen erilaisia fyysisiä liikkeitä kuin eurojen).

Tasojen välillä vallitsee siis *toteutumissuhde*: alempi fysikaalinen taso on tietyn ylemmän tason kuvauksen toteuma, mutta ylempi taso ei determinoi tai vaadi tiettyä täsmällistä toteumaa, vaan toteuma on alimääräytynyt (alideterminoitu) suhteessa ylemmän tason kuvaukseen. Tämänkaltaiset toteutumisrelaatiot ovat arkipäiväisiä: esimerkiksi Finlandia -sävelruno voi toteutua lukemattomilla tavoilla eri esittäjien toiminnoissa, mutta mikään tuossa taideteoksessa itsessään ei edellytä ketään tiettyä esittäjää. Toteutumissuhteen filosofiset ongelmat liittyvät pikemmin ylemmän tason kuvauksen ontologiaan (miten Finlandia tai talous tai semanttiset relaatiot oikeastaan ovat olemassa), mutta lienee joka tapauksessa selvää, että meillä ei ole mitään syytä pitää Finlandian, taloudellisten transaktioiden tai merkityssuhteiden olemassaoloa sinänsä mitenkään epäilyksenalaisena. Olennaista on huomata, että toteutumissuhteen osapuolina olevat eri tasojen kuvaukset ovat toisistaan riippumattomia. Ylemmän tason toteutuminen ei edellytä tiettyä yhtä täsmällistä toteutumisen tapaa, eikä alemmalla tasolla kuvatasta ilmiöstä välttämättä seuraa minkään tietyn ylemmän tason kuvauksen alaan kuulumista (esimerkiksi samat fysiologiset liikkeet voinee tehdä ilman, että toiminta on välttämättä taloudellisessa kuvauksessa vaihtorahan palauttamista). Koska eri tasoilla olevien kuvausten välillä ei ole suoraviivaista ja säännönmukaista vastaavuutta, olisi niiden sekoittaminen keskenään suuri virhe.⁸

⁸ Alaviitteessä 2 mainitsemani Peacocken näkemys on virheellinen tässä esitetyn ongelman vuoksi: hän sekoittaa kaksi erillistä kuvausten tasoa toisiinsa.

Kuten aiemmin totesin, Fodor käsittää mielen merkitysrelaatiot funktioksi sisäisestä (syntaktisesta) tilasta kontekstiin. Tietty sisäinen tila voi saada erilaisia merkitysrelaatioita erilaisissa konteksteissa, minkä vuoksi pelkkä sisäinen kuvaus ei ole tyhjentävä kuvaus mielen semanttisista ominaisuuksista. Vastaavasti tietyt merkitysrelaatiot voivat toteutua useilla erilaisilla sisäisillä tiloilla (esimerkiksi ajatuksellinen sisältö ”kuulemani sävelet ovat Finlandiasta” voi toteutua hyvin erilaisissa representaatiojärjestelmissä tai neurologisissa konfiguraatioissa). Uskomusten viittaussuhteet kohteisiinsa, sen paremmin kuin uskomusten totuuskaan, eivät palaudu *vain* niiden sisäisiin ominaisuuksiin. Fodorin ajatus on, että jotta funktiot mielen (syntaktisista) tiloista ulkoisiin asiantiloihin olisivat ylipäättään mahdollisia, on oltava tiettyjä *psykologisia lakeja*, joiden johdosta nuo funktiot pysyvät vakioina, ja ainoastaan ne ”alkiot” vaihtelevat, joihin funktio liittyy tietyn syntaktisen tilan eri konteksteissa. Juuri näiden lakien vakioisen toiminnan vuoksi Oskarin ja kaksois-Oskarin identtiset mielentilat voivat saada erilaiset viittauksen kohteet eri konteksteissa, meidän Maasamme ja kaksoismaassa. (Fodor 1994; 2008).⁹

Fodor haluaa tehdä psykologisista laeista selitysvoimaisia monien yleisemmän tason ilmiöiden suhteen. Erityisesti Fodor haluaa selittää sen, miten erilaiset aivotilat sekä erilaiset syntaktiset mielentilat voivat toteuttaa identtiset ajatukset. Tämä on Putnamin ongelman vastakohta. Putnam ihmetteli, miten kaksi identtistä aivo- ja mielentilaa voivat ilmentää eri ajatuksia. Fodor taas ajattelee, että mentaaliset prosessit voivat toteutua erilaisissa materiaalisissa konfiguraatioissa ja erilaisilla mielillä voisi olla tyystin erilaiset syntaksit. Niinpä yksi selitystä kaipaavista ilmiöistä on se, miten erilaisen alemman tason implementaatiot (eri aivotilat, eri syntaksit) voivat toteuttaa identtiset

⁹ Fodorin näkemys psykologisista laeista on joistakuista muistuttanut paljon Fregen näkemystä mielistä ja niiden roolista viittausten määrittämisestä, mutta Fodor on kiistänyt tämän (Fodor 2008, 86 alaviite 56).

ajatukset. ("Toteuttaa identtiset ajatukset" -ilmauksen voi ymmärtää tässä tarkoittavan esimerkiksi samaa kuin "toteuttaa saman suhteen samaan proposition".) Fodorin lähtökohta psykologisille laeille on siis se, että niiden selitysalaa ovat sellaiset ylemmän tason yhdenmukaisuudet ajattelussa (esim. viittaaminen ja merkitykset), jotka eivät ole palautettavissa alemman tason, mielen syntaksin tai aivotilan yhdenmukaisuuksiin. Tästä johtuen hänen selityksensä etenee ikään kuin rakentamalla tietyn laajemman selityksen mielen syntaktisen kuvauksen ympärille, joka selittää mielen semantiikan (Fodor 2008, 91). Tätä voidaan nimittää *konstruktivistiseksi näkemykseksi* tai *alhaalta ylös - näkemykseksi*.

Tässä on mahdollisesti merkittävin ero Millin ja Fodorin ajattelun välillä. Esimerkiksi Millin alkuperäinen intuitio näyttää liittyvän vain vähän jos lainkaan kysymykseen *synkroniasta* mielen ja maailman välillä – sen ytimessä on pikemmin ajatus, että mielen sisäinen prosessointi täytyy käsittää mielen ja maailman välisten kognitiivisten suhteiden toteutumisenä paljon vahvemmassa mielessä kuin pelkkänä synkroniana (Millin muotoilut omasta näkemyksestään ovat sen verran epämääräisiä, että tulkintani ei ole mitenkään kiistaton). Fodorin konstruktivismissa puolestaan tietty sisäinen prosessointi on lähtökohta, jonka vastaavuuden tai synkronian maailman kanssa ylemmän tason psykologiset lait selittävät. Millin filosofisesti merkittävä huomio, ainakin siten kuin minä häntä tulkiten, on se, että ajattelun sisäinen prosessi mahdollisine komputaatioineen ja syntaktisine ominaisuuksineen on pohjimmiltaan alisteista niille relaatioille, jotka tuo prosessi mahdollistaa kognitiivisen toimijan ja tämän ympäristön välille. Näin syntaktisia kuvauksia pitäisi aina tulkita kognitiivisten relaatioiden viitekehysten läpi. Fodorin näkemys on käytännössä päinvastainen – psykologiset lait liittävät erilaiset syntaktiset tilat maailmaan, ja tietyssä mielessä selittämisen suunta on mielen sisäisistä prosesseista maailmaan – siinä missä Millin selittämisen suunta on kognitiivisista suhteista sisäisiin prosesseihin. Millille kysymys on siitä perustavasta ajatuksesta, että ajattelumme on sisällöil-

tään ja luonteeltaan merkittävässä määrin mielen ulkoisen todellisuuden määräämää. Fodorin tähtäimessä ovat pikemmin vain tiettyjen kognitiivisten suhteiden toteutumisen ehdot tietyille syntaktiselle tilalle tietyssä kontekstissa.

Esimerkki sydämen toiminnan kuvaamisesta valaisee ongelmaa, joka nähdäkseni vaivaa Fodorin lähestymistapaa. Sydämen toimintaa voidaan kuvata pelkkänä sydämen sisäisenä prosessina, esimerkiksi selittämällä sinussolmukkeen toiminnan vaikutus sydämen sykkeeseen. Voimme tietysti laajentaa tätä kuvausta ja selittää sinussolmukkeen toimintaa viittaamalla lainomaisiin tapoihin, joilla esimerkiksi elimistön rasitus vaikuttaa sinussolmukkeeseen. Kuitenkin pelkkiin lainomaisiin suhteisiin viittaaminen ei näytä tuottavan riittävää ymmärrystä selitettävistä ilmiöistä. Viimekädessä sydämen toiminta tulee ymmärrettäväksi vasta, kun sen toiminta kuvataan tiettyjä tarkoitusperiä toteuttavaksi suhteessa elimistön verenkiertoon ja aineenvaihduntaan (ja näiden kriittinen merkitys koko organismin elämälle ja eloonjäämiselle huomioidaan). Mikä tahansa paikallinen ilmiö, kuten sinussolmukkeen rytmiä kiihdyttävä toiminta, näyttää siis *selityksellisesti alisteiselta* jollekin laajemmalle toiminnalle, kuten elimistön aineenvaihdunnalle, jota tuo paikallinen ilmiö osaltaan toteuttaa.

Fodorin näkemys ei kuitenkaan toimi edellisen esimerkin tavalla. Fodor pikemmin haluaa sulkeistaa mielen toiminnan itsenäisesti ymmärrettäväksi ja selitettäväksi alueekseen ja jättää psykologialle ulkoisen osan (viittaamisen, merkitykset) erillisten selitysten hoidettavaksi. Hän siis haluaa ikään kuin eristää sinussolmukkeen toiminnan omaksi ilmiökseen, jolla on sisäinen rakenne, jonka selittäminen itsenäisesti, ilman sydämen, aineenvaihdunnan ja organismin eloonjäämisen huomioimista on mahdollista sekä suotavaa. Näyttää siltä, että Fodor suuresti yliarvioi mahdollisuuksiamme ymmärtää ja selittää ajattelun sisäistä prosessia ottamatta huomioon sitä laajempaa ilmiötä, jonka osana nuo prosessit toimivat.

Psykologisten lakien avulla Fodor haluaa selittää, miten tiettyillä syntaktisilla rakenteilla päätyy olemaan tietyt suhteen ul-

koisiin olioihin ja asiointiloihin. Nähdäkseni luontevampi selityksen suunta olisi itseasiassa päinvastainen, hieman Millin näkemystä mukaillen. Voimme helposti nähdä päinvastaisen suhteen merkitysrelaatioita hallitsevien psykologisten lainalaisuuksien ja mielen syntaktisten tilojen välillä kehitellessämme vaikkapa evolutiivisia selityksiä mielen luonteelle. Miksi miellemme syntaktiset operaatiot ovat enemmän tällaisia kuin tuollaisia? Mahdollisesti juuri siksi, että miellemme on evolutiivisesti kehittynyt ympäristössä, jossa myös mainitut lainalaisuudet vaikuttavat. Toisin sanoen maailmassa vaikuttavat lainalaisuudet määräävät ajatteluamme tietynlaiseksi, ja tästä syystä mielenprosesseillamme on myös enemmän tämänlainen kuin tuollainen syntaktinen muoto. Samanlainen suhde vallitsee kielen syntaksin ja maailman välillä: esimerkiksi ainesanat (so. ainesta nimittävät sanat) taipuvat eri tavoin kuin nimisanat (asioita nimeävät sanat). Tavallisissa tapauksissa edellisillä ei ole monikkomuotoa, eivätkä lukusanat sovi niiden määreiksi (so. monia tuoleja vrt. *monia vesiä). Ilmiö ei ole selitettävissä pelkästään syntaksin itsenäisillä piirteillä, vaan se johtuu ymmärrettävästi siitä, että asiat, joista noiden sanojen avulla puhutaan, käyttäytyvät eri tavoilla. Vesi ei – ainakaan ilman astioita, kuten pulloja – muodosta laskettavia kokonaisuuksia, mutta tuolit ovat (tyypillisesti) jo itsessään riittävän erillisiä olioita tullakseen lasketuiksi. Psykologiset lainalaisuudet voidaankin käsittää samalla tavalla maailmassa vallitseviksi piirteiksi, jotka määräävät monia ajatteluamme liittyviä ilmiöitä ja näin tuottavat näennäisesti pelkille syntaktisille ominaisuuksille myös tietyt mielenulkoiset suhteet ulkomaailman olioihin. Molemmassa tapauksissa, kielessä ja mielessä, on siis kyse siitä, että ulkoiset (semanttiset) suhteet määräävät ainakin osan syntaksista. Näin ollen oikea selityssuunta olisi edetä mielen ja maailman välisistä suhteista tarkastelemaan syntaktisia operaatioita, eikä ensisijaistaa mielen prosessien syntaktista tarkastelua.¹⁰

¹⁰ Filosofian historiassa tämä sama kysymys on tullut esiin esimer-

Siten kuin minä sitä tulkitsen, Descartesin mielenfilosofia näyttää tekevän juuri jotakin tämänkaltaista. Descartesin näkemys sisältää osioita, jotka näyttävät minusta vahvasti yritykseltä asettaa mielen sisäinen prosessi alisteiseen asemaan suhteessa mielen ja maailman välisiin merkitysrelaatioihin. Tästä huolimatta Descartesin ja Fodorin teorioilla on monia yhtymäkohtia, jotka tekevät näiden kahden näkemyksen vertailusta mielekäästä.

Descartes ja kognitiivinen identiteetti

Descartesin mielenfilosofia on monella tapaa aristoteelisen filosofian jälkeläinen, vaikkakin Descartes suhtautuu perintöönsä usein hyvin kriittisesti. Tarkasteluni kannalta olennainen seikka koskee hänen käsitystään havainto- ja ajattelurelaatiosta identiteettirelaationa, joka on monella tapaa yritys säilyttää aristoteelinen käsitys *kognitiivisesta identiteetistä* sitoutumatta kuitenkaan aristoteeliseen muoto-materia-oppiin.

Monille aristoteelikoille, kuten Tuomas Akvinolaiselle, kognitio edellytti identiteettirelaatiota mielensisällön ja ajattelun kohteen välillä. Aristoteelinen metafysiikka mahdollisti tämän esittäessään olioiden muodostuvan kahdesta komponentista, materiaasta ja muodosta, eli *formasta*. Forma ”informoi” materiaa, joka itsessään käsitettiin homogeeniseksi ja indviduoitumattomaksi, ja forma oli näin vastuussa olioiden varsinaisesta identiteetistä. Saman forman oli mahdollista välittyä mieleen erillään materiaasta ja toimia näin myös tuota oliota koskevan ajatuksen konstituenttina.¹¹

kiksi 1800-luvun uuskantilaisten kiistoissa, nk. Trendelenburg–Fischer-kiistassa, jossa esittämäni vaihtoehto olisi jotakuinkin Trendelenburgin puolustama ”laiminlyöty vaihtoehto”. Kiitän Lauri Kalliota tämän keskustelun tuomisesta tietoisuuteeni hänen tutkimuksensa esittelyn yhteydessä.

¹¹ Aristoteleen käsitys ajattelusta identiteettirelaationa löytyy hänen *Sielusta* -teoksestaan (Aristoteles 2006). Akvinolaisen esitys ”formaalista identiteetistä” löytyy hänen kommentaaristaan em. teokseen (Aquinas, Foster, ja Humphries 1951).

Pelkän havainnollistuksen nimissä voidaan sanoa, että aristoteelisen näkemyksen mukaan ajatellessani tiettyä oliota ovat tuon olion representaation syntaktiset piirteet samalla kyseisen olion itsensä konstituivia syntaktisia piirteitä. Syntaktisuus ei tässä tarkoita tietenkään kielellisten symbolien yhdistelyä koskevia sääntöjä, vaan ehkä pikemmin olemassaolon yleisiä loogisia ominaisuuksia koskevia lainalaisuuksia. Ajatuksen keskeinen osa on identiteettirelaatio olion ja sen mielellisen representaation ominaisuuksien välillä, ei pelkkä satunnainen vastaavuus. Toisin sanoen sama asia, joka konstituoii olion ja vastaa sen identiteetistä, myös konstituoii ajatukseni tuosta oliosta ja vastaa ajatukseni identiteetistä (siis siitä, että ajatukseni koskee juuri tuota oliota). Descartesin voidaan tulkita pyrkivän säilyttämään näkemyksen ajattelusta *identiteettirelaationa* ajattelijan ja ajattelun kohteen välillä, mutta pyrkivän tekemään tämän ilman aristoteelikoiden käsitystä formista.

Descartesin yritys muotoilla ajattelurelaatio ajattelijan ja objektin väliseksi identiteettirelaatioksi perustuu hänen käsitykselleen *objektiivisesta realiteetista*.¹² *Realiteetti* yleisemmin otettuna on kvantitatiivinen määre, joka ilmentää olion olemassaolon määrää. Kaikella olemassa olevalla itsessään, fyysisistä objekteista loogisiin totuuksiin, on aina itsenäistä realiteettia. Tätä itsenäistä, intrinsistä realiteettia hän nimittää *formaaliksi realiteetiksi*. Descartes toteaa yksilöolioilla olevan enemmän realiteettia kuin niiden ominaisuuksilla, koska yksilöoliot ovat riippumattomampia muista olioista toisin kuin ominaisuudet, jotka ovat välttämättä jonkin olion ominaisuuksia. Juuri olemassaolon itsenäisyys näyttää olevan se kriteeri, jonka perusteella realiteetin määrä voidaan arvioida. Kaikkein eniten realiteettia on Jumalalla, joka on ainut todella mistään muusta riippumaton olio.

¹² Descartesin keskeiset esitykset objektiivisesta realiteetista löytyvät 3. mietiskelystä (2002, 49–57; CSM2, 27–36; AT7, 40–52) ja vastauksista ensimmäisiin vastaväitteisiin (ks. alla). Viitteen lyhenteet CSM ja AT viittaavat Descartesin töiden ”virallisiin” englanninkielisiin ja ranskan-latinankielisiin laitoksiin samassa järjestyksessä (ks. *Lähteet* alla).

Descartesin kausaatioteoria on erottamaton osa hänen näkemystään ajattelusta, mutta sen tarkasteleminen ei ole välttämättöntä tässä yhteydessä.

Descartesin mukaan ajattelu on luonteeltaan tietyn olion realiteetin tulemista mieleen. Olio, jolla on formaalia realiteettia, esimerkiksi aurinko, on kausaalisessa kytköksessä mieleen ja välittää sille oman realiteettinsa. Näin auringon ajatteleminen on Descartesin mukaan auringon realiteetin tulemista mieleen ja samalla auringon itsensä olemassaoloa mielessä. Koska oliot itsessään, siten kuin ne ovat olemassa itsenäisesti eli formaalisesti, eivät voi olla mielessä, kutsuu Descartes sitä olemassaolon tapaa, joka olioilla on ajattelijan mielessä, *objektiiviseksi olemassaoloksi* tai *realiteetiksi*.

Mutta sisältääkseen juuri jonkin tietyn objektiivisen reaalisuuden, tämän idean täytyy varmasti saada se joltakin syytä, jossa on ainakin saman verran formaalista reaalisuutta kuin ideassa on objektiivista. Jos nimittäin oletamme, että ideassa esiintyy jotain, mitä ei ole ollut sen syyssä, niin se on silloin saatu tyhjästä. Ja vaikka tuo olemisen modus, jolla olio on objektiivisesti ymmärryksessä idean kautta, olisi kuinka epätäydellinen, se on silti varmasti jotakin, eikä se siis voi olla tyhjästä peräisin. (Descartes 2002, 50; CSM2, 28–29; AT7, 41.)

Keskeistä Descartesin näkemyksessä mielenulkoisesta ja sisäisestä realiteetista on, että hänelle mielessä oleva objektiivinen realiteetti edellyttää aina mielenulkoista formaalia realiteettia. Itse asiassa Descartesin käsitys eroaa juuri tässä häntä edeltäneiden skolastikkojen näkemyksistä, vaikka objektiivisen olemassaolon käsitteistö muutoin onkin lainaa näiltä. Esimerkiksi Suárez, jonka filosofialle Descartes on skolastikoista ehkä eniten velkaa, esitti mielessä voivan olla myös objektiivista realiteettia, vaikka se ei olisikaan minkään mielenulkoisen olion realiteettia. Descartes kieltää tämän mahdollisuuden yksiselitteisesti.¹³

¹³ Olen tarkastellut Descartesin ja Suárezin yhteyttä väitöskirjassani

Vastauksissaan *ensimmäisiin vastaväitteisiin* Descartes esittää näkemyksen, jonka mukaan jopa fiktiivisten olioiden ajattelu koostuu pohjimmiltaan jostakin todellisesta peräisin olevan realiteetin manipuloinnista. Kun nerokas insinööri kuvittelee mielessään mutkikkaan koneen, ei tuokaan kuvitelma sisällä mitään, mikä ei olisi peräisin jostakin muusta. Koneen rattaiden ja jousten ideoiden realiteetti on peräisin jostakin havaitusta, ja ennennäkemätön järjestys osien välillä on realiteettinsa velkaa insinöörin mielen yhdistelykyvylle. Mieli itse ei Descartesin mukaan kykene tuottamaan mitään uutta ”sisältöä” eli objektiivista realiteettia. (Descartes 2002, 96–97; CSM2, 75; AT7, 103–104.)

Voimme kuvitella vaikkapa siipirasaluksen keksijän, joka kuvitteli siipirasaluksen ennen kuin ainuttakaan siipirasalusta oli olemassa. Nerokas keksijä tiesi jotain höyrykoneista ja laivoista, ja hänellä oli käsityksiä voimansiirrosta akseleiden avulla ja siivekkäistä rattaista, joita myllyissäkin käytetään vesipyörinä. Kuitenkaan hänen mielensä ei tuottanut mitään todella uutta kyseiseen ideaan. Ainoa sen todella ennennäkemätön piirre on osien kompositio: aiemmat ideat höyrykoneista, laivoista, voimansiirrosta ja vesipyöristä muodostavat idean sisällön, eikä mikään osa uutta järjestystä lukuun ottamatta ole ”luotu tyhjästä”.

Kertaan seuraavassa oman tulkintani Descartesista, mutta en argumentoi sen puolesta tässä, koska olen tehnyt sen muualla (Sinokki 2016, 2017a, 2018b, 2018a).

Ei ole erityisen kaukaa haettava nähdä samankaltaisuuksia Descartesin näkemyksen ja fodorilaisen komputationalismin välillä: Fodorin komputationalismi olettaa, että mielen operatiot tehdään olemassa olevia primitiivisiä komponentteja, käsitteitä, yhdistelemällä, eli että ajattelu on *kompositionaalista*. Mielen kyky tuottaa potentiaalisesti äärettömästi uusia yhdistelmiä – siis mielen *produktiivisuus* – on selitettävissä tämän kompositionaalisuuden avulla. Descartesin näkemys ajattelusta näyttää jakavan fodorilaisen komputationalismin kanssa keskeisesti

(Sinokki 2016, 102–13).

juuri käsityksen ajattelun kompositionaalisuudesta ja produktiivisuudesta.¹⁴

Fodor käsitti seuraavansa Descartesia siinä, että hän oletti mielen sisäisen ”leksikon”, eli syntaktisten muotojen perusvaraston olevan sisäsyntyinen ja rajallinen. Vaikka Descartes tosiaan esittää mielessä olevan synnynnäisiä ideoita, ilmentää tämänkin näkemys itseasiassa hänen käsitystään objektiivisesta realiteetista. Esimerkiksi loogiset tai ikuiset totuudet, kuten ristiriidan laki, ja monet peruskäsitteistä, kuten yksilön käsite, ovat kyllä synnynnäisiä, mutta niiden sisältämä objektiivinen realiteetti on silti jonkin mielen ulkopuolella formaalisesti olemassa olevan realiteettia. Descartes ei siis nähdäkseen hyväksyisi esimerkiksi ajatusta ristiriidan laista pelkkänä syntaktisena, muodollisena totuutena vailla semanttista sisältöä, vaan mielessäni oleva idea ristiriidan laista olisi hänelle täysin riippuvainen mieleni ulkopuolisesta ikuisesta totuudesta, siis siitä, että ristiriidan laki pätee koko luomakunnasta kategorisesti. Idea on toisaalta synnynnäinen juuri siksi, että tuo laki pätee kategorisesti, siis myös minusta ja mielestäni. Näin ollen kyseisen objektiivisen realiteetin tapauksessa minun ei tarvitse saada *a posteriori* tietoa mistään mieleni ulkopuolisesta, vaan voin ajatella ristiriidan lakia, koska voin löytää sen omasta mielestäni täysin riippumatta havainnoistani. Mielen ulkoisten yksilöolioiden ajatteluun tarvitaan kuitenkin lisäksi kausaalinen mekanismi, joka mahdollistaa olioiden realiteetin tulevan mieleen (esim. valo). Käytännössä Descartes ajattelee esimerkiksi aistihavainnon kausaalisten suhteiden tekevän juuri tämän. Hän myös todennäköisesti hyväksyisi Saul Kripken ajatuksen siitä, että erisnimien viittaus perustuu nimen käytön historiallis-kau-

¹⁴ Descartesin ajattelussa kompositionaalisuus tulee näkyviin myös muilla tavoin kuin vain objektiivisen realiteetin tapauksessa. Erityisesti hänen näkemyksensä synnynnäisistä ideoista ja ikuisista ja muuttumattomista luonnoista mukautuvat tähän malliin (ks. esim. Burgess 2018). On syytä huomata, että Fodor perustaa oman ”kartesiolaisuutensa” pikemminkin Descartesin innatismille kuin tämän objektiivista realiteettia koskevalle näkemykselle.

saaliselle välittymiselle kieliyhteisössä (ks. Kripke 1980): tiettyssä mielessä Descartesin näkemys objektiivisesta realiteetista on metafyyminen malli semanttisten suhteiden luonteesta hyvin yleisellä, ontologisella tasolla. Kripken näkemyksen voidaan ajatella ilmentävän, tosin vain historiallis-kausaalisen teorian osalta, tätä yleisempää mallia semanttisten suhteiden ontologiasta.

Keskeinen ongelma, joka Descartesin malliin liittyy, koskee hänen pyrkimystään muotoilla kognitiivinen relaatio ajattelijan ja objektin välillä identiteettirelaatioksi. Tämän hän pyrkii tekemään esittämällä mielessä olevan objektin, kuten auringon, olevan *sama* objekti kuin se, jolla on formaalista olemassaoloa mielen ulkopuolella (ks. yllä). Ero objektiivisen realiteetin ja formaalisen realiteetin välillä tähtää nimenomaan sen vahvistamiseen, että on vain yksi olio, jolla on kaksi olemassaolon muotoa, toinen mielen ulkopuolella, toinen mielessä. On helppo nähdä, että nämä erilliset "olemassaolon muodot" eivät kuitenkaan voi olla identtisiä. Aurinko taivaalla on olemassa fyysisenä plasma-pallona, mutta mielessäni aurinko on olemassa korkeintaan mentaalisenä ideana tai representaationa. Leibnizin lain perusteella identtisyydestä seuraa myös ominaisuuksien samuus, joten kyse ei voi olla identiteetistä. Kuten Paul Hoffman tästä vaikeudesta päättelee, näyttää Descartesin näkemys hänen yrityksestään huolimatta romahtavan näkemykseksi, jossa mielellinen idea on vain mielenulkoisen olion representaatio mielessä vailla todellista "kognitiivista identiteettiä" noiden kahden olion välillä.¹⁵ (Hoffman 2002.)

¹⁵ Hoffmanin johtopäätös vaikuttaa hieman hätäiseltä. Esimerkiksi David Wiggins (2001) on analysoinut identiteettirelaation logiikkaa siten, että identtisyys on aina suhteellista tiettyyn olioon tai substanssiin nähden. Esimerkiksi minä voin olla *sama ihminen* kuin eilenkin, vaikka monet ominaisuuteni materiastani vaatteisiin ovat vaihtuneet. En tietystikään ole sen enempää *sama materiakokonaisuus* kuin olen *samoihin vaatteisiin pukeutunut*. Koska taivalla oleva aurinko on *sama aurinko* kuin se *aurinko, joka on mielessäni*, ei ole ainakaan itsestään selvää, että Descartesin muotoilu rikkoo Leibnizin lakia. Näyttää itse asiassa siltä, että Hoffmanin tulkinta on riippuvainen tietyistä tulkinnallisista

Kognitiivinen identiteetti ja kuvausten tasot

On kuitenkin toinen tapa ymmärtää Descartesin hanke ajattelun mekanismien selittämiseksi. Voimme nimittäin asettaa Descartesin näkemyksen komponentit hierarkiaan, jonka perusteella ajatus kognitiivisesta identtisyydestä näyttää ainakin johdonmukaisemmalta kuin äskeisen ongelman valossa.

Yleisellä tai korkeimmalla tasolla kognitiossa on oletettavasti kyse tiedollisista suhteista kognitiivisen agentin ja tämän ajattelun kohteiden välillä. Esimerkiksi niin kutsutun arkipsykologian käsitteet ovat usein sovellettavissa tällä kuvauksen tasolla. Kyse on sellaisista kognitiivisista tiloista kuin tietämisestä, uskomisesta, epäilystä, tarkoittamisesta, ja niin edelleen. Teknisemmin ilmaistuna kyse on propositionaalisista asenteista, jotka edellyttävät suhdetta (tyypillisesti) mielestä erillisiin olioihin. Tällä tasolla kuvaus voi olla jotakin sen kaltaista kuin "JS ajattelee aurinkoa". "JS" nimeää minut, "aurinko" auringon, ja "ajattelee" nimeää tietyn kognitiivisen relaation, jossa siis tässä tapauksessa minä olen subjekti-*relatum*, ja aurinko on objekti-*relatum*. Tämän tason voidaan ajatella muodostavan ajattelun yleisimmän, ja hierarkiassa ylimmän, tason. Descartesin filosofiassa substanssit ja ominaisuudet kuuluvat tälle tasolle, ja tarkastelun kohteena ovat ne relaatiot, jotka vallitsevat substanssien välillä (esim. ajattelija-aurinko) tai substanssien ja toisten substanssien ominaisuuksien välillä (ajattelija-vahanpalan kiinteys).

Seuraavan alemman tason kuvaus on kuvaus mekanismista, joka toteuttaa tämän ylempään tason kuvauksen. Descartesin näkemys objektiivisesta realiteetista on luontevaa käsittää eräänlaisena metafyyssisenä mekanismina, joka toteuttaa ajattelu-relaation ajattelijan ja objektin välillä: voin ajatella aurinkoa vain, jos juuri aurinko itse on tullut mieleeni oikealla tavalla.¹⁶ Mikään muu kokoonpantu ideani, jonka objektiivinen

presuppositioista, joita olen nimittänyt *representationalistiseksi viinoumaksi* (engl. *representationalist bias*) (Sinokki 2016).

¹⁶ Attribuoin Descartesille tässä näkemyksen, jota Kripke puolustaa käsitellessään yksisarvisten ajattelemisen mahdollisuutta – emme voi

realiteetti on peräisin jostain muusta kuin auringosta, ei voi olla auringon idea, vaikka se piirteiltään olisi erottamaton auringon ideasta (ks. myös Descartes 2002, 48–49; CSM2 17; AT7, 39). Descartesin malli toisin sanoen kertoo meille, että tämä ylemmän tason kuvaus on tosi *jos ja vain jos* auringon realiteetti on JS:n mielessä objektiivisesti. Realiteettien ontologian taso kertoo siis myös millä riittävin ja välttämättömin ehdoin ylemmän tason relaatio toteutuu.

Voimme huomata, että kun erotamme kognitiivisten relaatioiden tason realiteettikuvauksen tasosta, myös Descartesin väite kognitiivisesta identiteetistä tulee ymmärrettäväksi. Kognitiivinen identiteetti voidaan käsittää ajatuksena, että edellinen ehto ”JS ajattelee aurinkoa, *jos ja vain jos* auringon realiteetti on JS:n mielessä objektiivisesti” viittaa ylemmän tason kuvauksen entiteetteihin (so. minuun, aurinkoon, ja ajattelu-relaatioon), ja siis edellyttää tiettyä kiinteää ontologista suhdetta minun mieleni ja auringon välillä. On tietysti selvää, että puhe ”realiteeteista” ei ole erityisen selkeää, mutta perusajatus on nähdäkseni hyvinkin ymmärrettävä: kuvausten tasojen hierarkia ikään kuin vakioi käytettyjen ilmausten viittaukset ylemmällä kuvauksen tasolla ja edellyttää samaa vakioista viittausta alemman tason kuvausten sisällöiltä.

On syytä täsmentää, mitä tarkoitan ”viittaussuhteiden vakioisuudella kuvausten välillä”.¹⁷ Kuten aiemmin totesin, kuvaukset ovat toisistaan riippumattomia: ylemmän tason kuvaus ei edellytä mitään tiettyä toteutumista, vaikkakin jotain – tätä tai tuota – toteutumisen tapaa se kysyy. Tästä erillisyydestä johtuen kuvausten välillä ei ole myöskään mitään sellaista suoraviivaista vastaavuutta, joka yhdistäisi eri kuvauksien sisällöt toisiinsa ja nivoisi ne yhteen yhdeksi kokonaiskuvaksi. Eräs perinteinen filosofinen dilemma näyttää olevan suoraa seurausta tästä. Aiemmin mainitsemani Locken kannattaman näkemyksen mukaan ylemmällä, kognitiivisten relaatioiden kuvausten

ajatella yksisarvisia, koska niitä ei ole ollut olemassa (Kripke 1980, 156–58).

¹⁷ Kiitän toista arvioijaa tämän täsmennystarpeen osoittamisesta.

(esimerkiksi tieto, havainto ja uskominen) tasolla kuvaus "auringon ajattelemisen" määräytyy ajattelijan ja auringon relaation perusteella (toisin sanoen, kuvatun ajatuksen identiteetti määräytyy osin sen perusteella, mikä olio aurinko sattuu olemaan), mutta mielensisäisen kuvauksen tasolla tuo relaatio ei ole käytettävissä. Syy ilmiölle on sama kuin komputationalismissakin: tämän kaltaiset mielenulkoiset relaatiot eivät ole osa *tuota* kuvausta, joten niihin vetoaminen mielen mekanismin selittämisessä ei ole mahdollista. Tästä näkökulmasta katsottuna Locken ajautuminen tyystin epätyytyttävään kuvaus plus tyydyttämisehdot -näkemykseen ideoiden representoimisen mekanismista vaikuttaa ymmärrettävältä. Toisaalta kuvausten tasojen sekaantumisen välttämiseksi on maksettava kova hinta: vaikka ylemmällä kuvauksen tasolla voimme väittää, että havaitessamme auringon me myös esimerkiksi *tiedämme* auringon (siis jotakin auringosta), ei alemmalla mielen kuvauksen tasolla vastaavaa väitettä voida oikeuttaa. Tämä johtuu siitä, että alemmalla tasolla oikeuttamisen on edettävä käytettävissä olevista resursseista lähtien, eivätkä mieli-maailma-suhteet kuulu niihin.¹⁸ Keskeistä on huomata, että meidän on tehtävä valintoja kuvausten tasojen asettamisessa hierarkiaan, joka muodostaa laajemman käsityksen: jos etenemme alhaalta ylös (kuten Fodor), joudumme selittämään toteutumisrelaation hyvin eri tavalla kuin ylhäältä alas edetessä (kuten Descartes). Descartesin näkemyksessä esimerkiksi kognitiiviselle relaatiolle, joka vallitsee JS:n ja auringon välillä, löytyy luontevat toteumat tarkasteltujen olioiden toiminnoista eri tasoilla kuvattuna. Kun siirymme kuvauksen tasolta toiselle, pysyvät keskeiset ilmaukset viittaukseltaan vakioina. Mihin ikinä ilmaus "aurinko" ylimmällä tasolla viittaakaan, viittaa se samaan olioon tai asiaan myös alemmilla tasoilla, koska alemmat taso ovat vain tuo

¹⁸ Tässä kuvattu ilmiö on tuttu keskusteluista, jotka koskevat internalistisen justifikaation ja ulkomaailman olemassaoloa koskevan skeptismin suhdetta. Tämän kiinnostavan aiheen tarkastelu ei tilan puutteen vuoksi ole mahdollista tässä kirjoituksessa.

ylimmän, ilmausten viittaukset kiinnittävän tason partikulaarinen toteuma. Esiteoreettisesti ajattelemme suhteita eri kuvausten tasojen välillä juuri tällä tavalla: koska vaikkapa tietty havaintorelaatio on aina relaatio, jolla on tietty partikulaarinen fyysikaalinen ja aivotilallinen toteuma, emme ehkä kiinnitä huomiota siihen, että teoreettisesti asiaa tarkastellessamme nuo kuvausten tasot on parempi pitää erillään, ja ne pitää erikseen nivoa yhteen. (Toisin sanoen hyvin erilaiset kuvaukset alemmalla tasolla voivat olla yhden ja saman ylemmän tason kuvauksen toteumia, ainakin teoreettisten mahdollisuuksien puitteissa: minun mielentilani tai fyysinen konstituutiosi olisi voinut olla kyseisessä tilanteessa erilainen, ja silti ylemmän tason relaatio "JS:n auringon ajattelemisen" olisi toteutunut.) Fodorin alhaalta-ylös-mallissa luonnollisen ennakko-oletuksemme ongelmallisuus tulee hyvin esiin: hänen täytyy kyetä selittämään, miksi tietty kuvaus mielentiloista todella on tietyn ylemmän tason kuvauksen toteuma tällä alemmalla tasolla. Juuri tämä selittää psykologisten lakien postuloimisen tarpeen Fodorin näkemyksessä, samoin kuin sen ongelman syntyminen, josta tuo tarve kumpuaa.

Descartesin näkemyksessä realiteettien tasoa alempi seuraava taso kuvaa mielensisäisesti niitä tiloja ja operaatioita, joita JS:n mieli edellä tarkastellussa tapauksessa tekee. Descartesille tämä on ideoiden ja ymmärryksen sekä arvostelman tekemisen taso; Fodorille tämä taso olisi komputationaalinen kuvaus mielen operaatioista. Tämä taso lisää oman panoksensa ylimmän tason kuvaaman kognitiivisen relaation riittäviin ja välttämättömiin ehtoihin. Tämän tason toteuma auringon ajattelemisesta saattaisi edellyttää esimerkiksi *singulaarisen olion idean* tai *käsitteen* esiintymistä JS:n mielen operaatioissa. Descartes ajattelisi tämän olevan substanssin synnynnäinen idea (Descartes 2002, 51; CSM2, 29–30; AT7, 42–45). Fodor taas kutsuisi samaista "ajattelun synnynnäistä atomia" pikemmin yksilöobjektin käsitteeksi kuin substanssin ideaksi (esim. Fodor 2008, 101 eteenpäin). Jälleen alempi taso on ylemmän tason kuvauksen paikallinen toteuma JS:n mielessä. Kaikki ne viittaussuhteet JS:ään ja aurinkoon, jotka esiintyvät ylimmällä tasolla, ovat yhä

toiminnassa tällä mielen sisäisen prosessin tasolla, vaikka tämän tason kuvaus sisältääkin itsessään viittauksia vain mielen ideoihin ja arvostelmiin tai komputaatioihin ja mielellisten representaatioiden syntaktisiin piirteisiin. Descartesin ja Fodorin mallit yhdistäen voitaisiin sanoa, että tietty representaatio ja sillä tehtävät komputaatiot konstituivat *auringon objektiivisen olemassaolon JS:n mielessä* jos ja vain jos mainittu ylemmän tason ehto täyttyy (eli aurinko ylemmän tason kuvauksessa on oikeassa suhteessa tiettyyn syntaktiseen representaatioon). Samoin nuo representaatiot ja operaatiot *konstituivat* sen sisäisen prosessin, joka on välttämätön ylimmän tason kognitiivisen relation, JS:n auringon ajattelemisen, toteutumiseksi.

Joudun ohittamaan ajattelun fysikaalisen implementaation tason tässä kirjoituksessa, vaikkakin sen tarkastelu olisi monella tavalla filosofisesti kiinnostavaa. Descartesin näkemyksessä juuri tällä tasolla esiin nousevat kysymykset ajattelun ja ruumiin (aivojen) erilaisista luonnoista ja niiden välisistä kausaalisista suhteista, jotka ovat tunnetusti ongelmallisia. Komputationalismissa monet kysymykset siitä, voiko ajattelu toteutua samanlaisena (identtisena), vaikka substraatit, joissa ajattelu toteutuu, olisivat hyvin erilaisia (aivokudos *vs.* logiikkapiirit), nousevat akuutisti esiin juuri tällä tasolla. Näiden sinänsä kiinnostavien kysymysten sijaan siirryn nyt tarkastelemaan, miten Descartesin objektiivinen realiteetti voi auttaa yhdistämään eksternalistien intuitiot komputationalistiseen mielen malliin.

Descartes ja komputationalismi

Voimme nyt käyttää käsitystä Descartesin objektiivisesta realiteetista paikkaamaan ongelmaa, jonka aiemmin havaitsimme vaivaavan komputationalismin ja eksternalismin yhdistämistä. Ongelma oli se, että mielentilojen tai representaatioiden relaatiot niihin olioihin, joihin ne viittaavat tai joita ne esittävät, eivät näytä olevan osa noiden mielentilojen tai representaatioiden syntaktista sisältöä. Koska vain mielen prosessien syntaktiset ominaisuudet ovat olennaisia mielen toiminnan kannalta,

eivät nuo ulkoiset suhteet näytä voivan vaikuttaa mielen toimintaan lainkaan, eikä niillä näytä voivan olla sellaista merkitystä ajattelulle kuin eksternalistit väittävät niillä olevan. Yksi ratkaisuehdotus tähän oli jättää mielenulkoiset merkitysrelaatiot mielestä erilliselle tasolle, jolla vaikuttavien psykologisten lakien perusteella tietyt representaatiot "suuntaavat" tiettyssä kontekstissa tiettyyn olioon ennemmin kuin johonkin toiseen. Tämä Fodorin kannattama ajatus on tiettyssä mielessä vain nykyaikaisempi versio kuvaus *plus* tyydyttymisehdot -mallista, jota Locke, ja mahdollisesti myös Hobbes, kannatti.

Esitetyn kuvausten tasojen hierarkian puitteissa ja Descartesilta lainaten voimme kuitenkin lähteä selityksessämme liikkeelle vastakkaisesta suunnasta kuin Fodorin konstruktivistinen alhaalta-ylös-malli. Descartesin kognitiivisen identtisuuden voidaan ajatella muodostuvan siitä, että kaikki relevantit alempien tasojen kuvaukset viittaavat vakioisesti ylimmällä tasolla ilmaistuun ajattelun objektiin, esimerkkitapauksessamme aurinkoon. Tämän hierarkian puitteissa ei ole mahdollista, että alemman, komputaationaalisen tason representaatio ei olisi auringon representaatio, koska malli käsitteellistää relaatiot siten, että olennainen representaatio on välttämättä osa ylimmällä tasolla määritellyn relaation toteumaa. Alempien tasojen kuvaukset ovat alisteisia siis niille viittaussuhteille, jotka esiintyvät ylemmillä tasoilla. Näin merkitysrelaatiot – kuten Descartesin objektiivinen realiteettikin – eivät ole yhdellä omalla kuvauksen tasollaan, vaan toimivat pikemmin eri kuvaukset toisiinsa liittävinä suhteina tasojen välillä. Esimerkiksi syntaksin tasolla ei voida puhua semanttisista relaatioista itsessään, vain ainoastaan suhteessa siihen laajempaan kuvausten hierarkiaan, jonka osa kyseinen syntaktinen kuvaus on (samoin kuin sinussolmukkeen rytmisäätelyä voidaan selittää ja ymmärtää vasta ottamalla huomioon sen merkitys aineenvaihdunnalle ja koko organismin elämän ylläpitämiselle). Ylin kuvaus siis *määrää* (so. determinoi) alempien tasojen käsitteellisen sisällön. Koska merkitysrelaatiot tulevat esiin vasta laajemmassa kuvausten hierarkiassa, ei ole ongelmallista, että pelkkiä syntaktisia ominaisuuksia omaavien representaatioiden tasolla ei voida käsitellä

semanttisia ominaisuuksia. Vaikka syntaktinen kuvaus mielen operaatioista näyttää aivan liian niukalta vangitakseen koko totuutta ajattelustamme, se voi olla riittävä kuvaus ajattelun *sisäisistä prosesseista*, kun se yhdistetään tuohon laajempaan selittävään kuvausten hierarkiaan, jonka huomioon ottaminen on välttämätöntä koko ajattelun luonteen ymmärtämiselle.

Tämä ”kartesiolainen komputationalistinen” malli siis yhtäältä selittää, miksi mielen komputaatioiden ja maailman välillä on lainomainen vastaavuus: komputaatioiden funktio ajattelussa on prosessoida informaatiota objekteista, tuottaa ennustuksia ja päätelmiä niiden käyttäytymisestä ja niin edelleen. Ne ovat samalla kuvauksia mekanismista, joka toteuttaa tietyn ylemmän tason kognitiivisen relaation. Toisaalta malli on ylhäältä alas etenevän hierarkian vuoksi riittävä varmistamaan, että alemman tason kuvauksen puitteissa pelkkiä syntaktisia ominaisuuksia omaavilla mielentiloilla tai representaatioilla on ylemmillä tasoilla sellaisia relaatioita maailmaan, jotka selittävät esimerkiksi referenssin ja siihen liittyvät, internalismille ja komputationalismille ongelmia tuottavat ilmiöt. Erillistä silta-periaatetta, joka kytkee syntaktisen kuvauksen kontekstiin ja selittää viittaamisen, ei tarvita tämän hierarkian vuoksi lainkaan.

Kysymys mielentilojen *per se* internalistisuudesta tai eksternalistisuudesta onkin tässä esitetyn mallin näkökulmasta sekaantunut, koska komputationaalinen kuvaus on lähtökohtaisesti vain osakuvaus niistä ilmiöistä, jotka (ylemmän kuvauksen tasolla) konstituivat ajattelijan ja maailman välisen kognitiivisen suhteen, ja joiden (yhtä mahdollista, eli todellisuudessa aktualisoitunutta) toteumaa komputationalistinen kuvaus kuvaa. Vaikeus eksternalismin ja komputationalismin yhteensovittamisessa johtuu siis siitä, että toteutumissuhteessa olevia ilmiöitä yritetään pakottaa yhden ja saman kuvauksen osiksi, jolloin sekaannus on vääjäämätön.

Lopuksi

Samankaltaisuuksistaan huolimatta Descartesin ja Fodorin näkemykset eroavat selittämisen tavoissaan merkittävästi. Siinä missä Descartesin lähestymistapaa voisi luonnehtia ylhäältä alaspäin eteneväksi analyysiksi ajattelun kognitiivisista relaatioista, näyttää Fodor tavoittelevan konstruktionistista selitystä, jossa kuvaus ajattelusta rakentuu ikään kuin mielen sisäisen prosessin kuvauksen upottamisesta laajempaan kontekstiin. Tästä selittämisen suunnassa olevasta erosta johtuen Fodor joutuu selittämään ajattelun merkitysrelaatiot ”kuvaus plus tyydyttymisehdot” -mallia muistuttavalla oletuksella psykologisista laeista, jotka varmistavat, että tietty sisäinen komputaationaalinen tila myös todella kytkeytyy tiettyihin olioihin tiettyssä kontekstissa. Descartesin näkemys sen sijaan vakioi analyysissä käytettyjen ilmausten viittaukset jo ensimmäisen, ylimmän tason kuvauksessa, jolloin alemmilla tasoilla, jotka ovat ylemmän tason toteumia (kuten mielen prosessin syntaktisen kuvauksen tasolla), ei voi käydä niin, että tietty syntaktinen tila esiintyisi ilman edeltävää kiinnittymistä tai viittaamista tiettyihin olioihin. Siinä missä Fodorin ongelma on kertoa, miten tietyt syntaktiset tilat saavat viittauksen mielenulkoisiin olioihin ja asiointiloihin, ei Descartesin tarvitse kiinnittää tähän suuresti huomiota, koska teoretisoidessamme vaikkapa mielessäni olevasta auringon ideasta työskentelemme valmiiksi sellaisen representaation kanssa, joka on syntyhistoriansa vuoksi kiinnitetty aurinkoon: se on JS:n ja auringon välisen kognitiivisen relaation toteuma.

On tärkeää huomata, että tässä kirjoituksessa on tarkasteltu vain mielen ”mekanismien” kuvausta, joka on mielen metafysiikan alaa. Esittämäni tulkinta (jota en tässä tekstissä edes pyrkinyt puolustamaan historiallisesti täsmällisenä tulkintana) on nähdäkseni hyvin hedelmällinen tapa tarkastella mielen sisäisiä prosesseja yhdistettynä mielen merkitysrelationaaliin ominaisuuksiin. Toisaalta lähes kaikki epistemologiset kysymykset ovat jääneet tarkasteluni ulkopuolelle. Miten mielen mekanismien kuvausten tasojen ymmärtäminen auttaa *minua*, *tiedollista toimijaa*, luovimaan epävarmojen uskomusten ja päätelmien

keskellä? Ehkä ainakin siten, että kun ymmärrän ajatteluni osana suurempaa kokonaisuutta – maailmaa ja sen kausaalisuhteita, ajattelun ja sen kohteiden välisiä suhteita ja ajatteluni sisäistä mekaniikkaa – voin huomata, että aivan kaikki tietokykyäni koskevat ongelmat eivät enää näytä uskottavilta tai edes johdonmukaisesti ylläpidettäviltä. Voin huomata, että ehkä tietokykyäni sittenkin riittää tiettyjen skeptisten vaihtoehtojen poissulkemiseen, edellyttäen, että olen riittävän hyvin ymmärtänyt sen laajan viitekehyksen, jossa tietokykyäni käytän. Kuten Descartes kirjoittaa:

[M]inun ei enää pidä kantaa huolta siitä, onko aistieni jokapäiväinen anti epätotta, vaan viime päivien liioitellut epäilykset on hylättävä naurettavina. Tämä koskee varsinkin perimmäistä epäilystä unesta, jota en erottanut valvetilasta. Nyt nimittäin huomaan, että niiden välillä on se suuri ero, että muisti ei milloinkaan yhdistä unia kaikkiin muihin elämän toimintoihin, kuten se yhdistää kaiken mikä valvovalle tapahtuu. (Descartes 2002, 84; CSM2, 61–62; AT7, 89.)

Pelkän muistikuvien koherenssin sijaan Descartes vaikuttaa viittaavan siihen, että hänen *Ensimmäisessä mietiskelyssä* esittämänsä skeptiset haasteet on mahdollista esittää vain tietyssä abstraktissa kontekstissa, mutta kun ymmärrämme oman olemassaolomme osana maailmaa ja ymmärrämme ajattelumme luonteen oikein, ei näitä haasteita enää ole mahdollista muotoilla koherentisti ja uskottavasti. Jos ideani auringosta on olemukseltaan juuri minun ja auringon välisen kognitiivisen suhteen toteuttaja (pikemmin kuin esimerkiksi representaatio, jolla on tietyt tyydyttymisehdot), on vaikeaa enää palata abstraktimassa kontekstissa esitettyyn huoleen siitä, että ehkä ideani ei vastaakaan mitään mieleni ulkopuolista. Vaikka en tässä tilassa voisi saada episteemistä varmuutta siitä, että ideani todella on tarkka representaatio oliosta, jonka idea se on, en kuitenkaan voisi epäillä enää sitä, etteikö idea olisi sentään jonkun mieleni ja sen ulkopuolisen maailman välisen suhteen toteuma. Voin siis erehtyä siitä, mitä oliota ideani esittävät ja siitä, esittä-

vätkö ideani nuo oliot oikein, mutta sitä, että ideat ovat yhteyteni johonkin mieleni ulkopuoliseen eivät ensimmäisen meditaation haasteet enää voi horjuttaa.

Suurempi ajatus, jota olen tässä kirjoituksessa kierrellyt, ansaitsee tulla artikuloiduksi: ajattelun ymmärtäminen yksityiskohtaisen mekaniikan tai laajojen relaatioiden tasolla on välttämättä riippuvaista siitä laajemmasta teoreettisesta ymmärryksestä, jota meillä ajattelun luonteesta on. Toisinaan juuri komputationalismin kehittäjät ovat kärsineet tietystä herkkyyden puutteesta filosofisen ymmärryksen merkitystä kohtaan. Mainitsen yhden kuuluisan esimerkin. Se on Alan Turingin teesi, että meidän ei pitäisi kysyä, voiko kone ajatella, vaan pikemmin voiko kone läpäistä niin sanotun Turingin testin – siis huijata ihmiskeskustelukumppania luulemaan itseään ihmiseksi (Turing 1950). Kysymys siitä, voiko kone huijata ihmistä, on kieltämättä mielenkiintoinen, mutta se on myös selvästi täysin yhteismitaton ehdotetun kysymyksen kanssa eikä korvaa sitä. Saadessamme vastauksen Turingin kysymykseen emme tietystikään olisi oppineet mitään ajattelun luonteesta tai koneajattelun mahdollisuudesta. Turingin pessimistinen suhtautuminen korvattavaksi ehdotettuun filosofiseen kysymykseen ei selvästikään tunnista, että ymmärryksemme ajattelusta voi kasvaa myös perinteisen metafysiikan tutkimuksen keinoin ja olennaisesti myös tavoilla, jotka eivät sulje pois Turingin ehdottamaa ”tieteellisempää” ajattelun kuvaamista. Päinvastoin, laajemmassa mittakaavassa molemmat kuvaukset voivat olla korvaamaton osa yhtä ja samaa hanketta.

Toisaalta monet perinteisen komputationalismin haastajat, kuten robotiikassa kannatettu *situationismi*, kaihtavat ajatusta, että ajattelu olisi vain sellainen sisäinen representaatioiden manipuloinnin prosessi, jota se komputationalismin mukaan on. Slogan ”maailma on paras itsensä representaatio” kuvaa hyvin tätä situationismin eetosta. Uskoakseni tässä kirjoituksessa esitetty malli antaa hyvän syyn ajatella, että jännitteisyys komputationalismin ja mielen ja maailman vuorovaikutuksellisuutta korostavien näkemysten välillä voi olla osin vain näennäistä.

Ajattelun täydellinen kuvaus, joka selittää myös meille keskeiset kognitiiviset relaatiot objekteihin kertomalla täsmällisen tarinan niiden mentaalisen ja fyysikaalisen toteutumisen tavoista, saattaa hyvinkin edellyttää, että sisäisiä prosessejamme kuvataan komputaatioina – ainakin, jos ajattelun systemaattisuutta halutaan kuvata täsmällisesti ja ”tieteellisesti”. Tämä ei silti tarkoita, etteikö mainittu slogan voisi olla totta noista prosesseista ainakin siinä mielessä, missä Descartes käsitti ulkoisten olioitten olemassaolon osaltaan konstituoivan mielensisältöjä, jotka puolestaan olivat keskeinen osa ajattelevan subjektin ja ajattelun kohteiden välisen suhteen toteutumista maailmassa. Maailma itse *on* noissa representaatioissa, koska ne ovat maailmasuhteemme toteutuma.¹⁹

Turun yliopisto

Kirjallisuus

- Tuomas Akvinolainen, Kenelm Foster ja Sylvester Humphries (1951), *Sentencia Libri De Anima - Commentary on Aristotle's De Anima*, New Haven: Yale University Press.
- Aristoteles (2006), *Teokset V*, toim. K. Näätäsaari ja J. Sihvola, Helsinki: Gaudeamus.
- Boghossian, Paul (1997), ”What the Externalist Can Know a Priori.”, *Proceedings of the Aristotelian Society* 97, 161–75.
- Burge, Tyler (1979), ”Individualism and the Mental.”, *Midwest Studies in Philosophy* 4, 73–121.
- Burgess, Steven (2018), ”Descartes’ Atomism of Thought: A Solution to the Puzzle about True and Immutable Natures.”, *Res Cogitans* 13:2.
- Descartes, René (1983), *Oeuvres De Descartes* 1–11, toim. C. Adam ja P. Tannery, Paris: Vrin. Lyhennetty viitatessa AT, jonka jälkeen ilmoitettu osa ja sivunumerot.

¹⁹ Kirjoituksen aiempi, lyhyempi versio esitettiin Suomen Filosofisen Yhdistyksen vuosikollokviossa ”Tekoäly, ihminen ja yhteiskunta” tammikuussa 2019. Haluan kiittää erityisesti Jani Hakkarasta hyödyllisistä kommentteista tuohon esitykseen. Kiitän Suomen Kulttuuri-rahastoa tutkimukseni pitkäjänteisestä rahoittamisesta.

- Descartes, René (1988), *The Philosophical Writings of Descartes* 1–3, käänt. J. Cottingham, R. Stoothoff ja D. Murdoch, osa 3 käänt. myös A. Kenny, Cambridge: Cambridge University Press. Lyhennetty viitatessa CSM, jonka jälkeen ilmoitettu osa ja sivunumerot; kolmas osa lyhennetty CSMK.
- Descartes, René (2002), *René Descartes: Teokset II*, toim. T. Aho ja M. Yrjönsuuri, Helsinki: Gaudeamus.
- Egan, Frances (1999), "In Defence of Narrow Mindedness.", *Mind and Language* 14, 177–94.
- Fodor, Jerry A (1987), *Psychosemantics*, Cambridge (Mass.): MIT Press.
- (1994), *The Elm and the Expert: Mentalese and Its Semantics*, Cambridge (Mass.): MIT Press.
- (2008), *LOT 2: The Language of Thought Revisited*, Oxford: Clarendon Press.
- Hobbes, Thomas (1651), *Leviathan or the matter, forme and power of a commonwealth ecclesiasticall and civill. (Leviathan, eli, Kirkollisen ja valtiollisen yhteiskunnan aines, muoto ja valta*, suom. T. Aho, Tampere: Vastapaino, 1999.)
- (1655), *Elementorum philosophiae sectio prima De corpore*. (Englanninkielinen käännös julkaistiin 1656; ensimmäisestä osasta on saatavilla moderni kriittinen englanninnos: *Computatio sive logica: Logic*, latinankielinen teksti, käännös ja kommentaari: A. Martinich; toim. ja johdantoossee I. C. Hungerland ja G. R. Vick, New York: Abaris Books, 1981.)
- Hoffman, Paul (2002), "Direct Realism, Intentionality, and the Objective Being of Ideas.", *Pacific Philosophical Quarterly* 83 (2), 163–79.
- Horowitz, Amir (2001), "Contents Just Are in the Head.", *Erkenntnis* 54 (3), 321–44.
- Kripke, Saul A. (1980), *Naming and Necessity*, Cambridge (Mass.): Harvard University Press.
- Loar, Brian (1988), "Social Content and Psychological Content." teoksessa R. H. Grimm ja D. D. Merrill (toim.), *Contents of Thought*, Tucson: University of Arizona Press.
- Locke, John (1689/1975), *An Essay Concerning Human Understanding*, toim. P. H. Nidditch, Oxford: Clarendon Press.
- Mill, John Stuart (1843), *A System of Logic, Ratiocinative and Inductive*, London: John W. Parker.
- Peacocke, Christopher (1994), "Content, Computation, and Externalism.", *Mind and Language* 9, 303–35.

- Peacocke, Christopher (1999), "Computation as Involving Content: A Response to Egan.", *Mind and Language* 14 (2), 195–202.
- Piccinini, Gualtiero (2004), "Functionalism, Computationalism, and Mental Contents.", *Canadian Journal of Philosophy* 34 (3), 375–410.
- (2008), "Computation without Representation.", *Philosophical Studies* 137 (2), 205–41.
- Putnam, Hilary (1975), "The Meaning of 'Meaning'", *Minnesota Studies in the Philosophy of Science* 7, 131–93.
- Sinokki, Jani (2016), *Descartes' Metaphysics of Thinking*, Turku: Turun yliopisto.
- (2017a) "Aistimukset ja interaktionismi: aistimukset mielen ja ruumiin yhteen nivojina Descartesin ontologiassa.", *niin & näin* 3, 91–98.
- (2017b), "Rasismi ja moraalinen sokeus." teoksessa J. Sinokki (toim.), *Rasismi ja filosofia*, Turku: Eetos ry, 37–58.
- (2018a), "Aistimusten merkitys havainnossa Descartesilla.", *Ajatus* 75, 151–80.
- (2018b), "Descartes ja havainto." teoksessa H. Laiho ja P. Remes (toim.), *Havainto*. Suomen Filosofinen Yhdistys, 53–62.
- Turing, Alan (1950), "Computing Machinery and Intelligence.", *Mind* 59 (236), 433–60.
- Unger, Peter (1984), *Philosophical Relativity*, Minnesota: University of Minnesota Press.
- Wiggins, David (2001), *Sameness and Substance Renewed*, Cambridge: Cambridge University Press.