

Robert Nozickin kokemuskone: Kahden todellisuuden teorioista ja niiden luomisesta

Samuel Kastell

Kandidaatintutkielma

Filosofian, poliittisen historian ja valtio-opin laitos

Filosofia

Turun yliopisto

Joulukuu 2024

Turun yliopiston laatuajärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

TURUN YLIOPISTO

Filosofian, poliittisen historian ja valtio-opin laitos / Yhteiskuntatieteellinen tiedekunta

KASTELL, SAMUEL: Robert Nozickin kokemuskone: Kahden todellisuuden teorioista ja niiden luomisesta

Kandidaatintutkielma, 25 s.

Filosofia

Joulukuu 2024

Tämä kandidaatintutkielma käsittelee Robert Nozickin (1938–2002) kokemuskonetta filosofisten teorioiden ja hypoteesien muodostamisen näkökulmasta. Tutkielman varsinainen tutkimuskysymys on se, miten filosofian alalla etenkin utilitarismin ja hedonismin kritiikistään hyvin tunnettu ajatuskoe, kokemuskone, voisi edesauttaa filosofisten teorioiden ja hypoteesien muodostamista.

Tutkielmani näkökulma perustuu Sara Kier Präemin ja Asbjørn Steglich-Petersenin artikkeliin ”Philosophical thought experiments as heuristics for theory discovery”, jonka avulla tarkastelen kokemuskonetta käsittelevää kirjallisuutta. Ajatuskokeet voivat tarjota vastaesimerkin, joka voi mahdollistaa jo olemassa olevan teorian kehittämisen ajatuskokeen avulla. Vastaesimerkit voivat puuttua jonkin teorian tai sen osa-alueen riittävyteen tai välttämättömyyteen. Näitä teorian parantamisia ajatuskokeen avulla Präem ja Steglich-Petersen nimittävät ad hoc -korjauksiksi. Ajatuskokeet voivat myös tarjota valmiin mallin, jota voi hyödyntää muunlaisiin filosofisiin tarkasteluihin etenkin muokkaamalla alkuperäistä ajatuskoetta. Metodologisesti kiinnostavia tapauksia ovat ajatuskokeet, jotka jonkinasteisen säännöllisyyden puitteissa auttavat uusien hypoteesien ja teorioiden muodostamista.

Tutkielmassani päädyn siihen johtopäätöksen, että kokemuskone on metodologisesti kiinnostava ajatuskoe, jota on käytetty ja on mahdollista käyttää teorianmuodostuksen apuna. Sitä voi hyödyntää sekä vastaesimerkkinä että mallina teorioiden ja hypoteesien muodostamisessa. Teorian kehittämisestä vastaesimerkin avulla esittelen Fred Feldmanin totuuden mukautettua asenteellista hedonismia, joka on teoriana vastaus valheellisten nautintojen ongelmaan, jonka tunnetuin esimerkki kokemuskone on.

Teorianmuodostukseen kokemuskone tarjoaa myös valmiin mallin, jota voi hyödyntää muunlaisiin filosofisiin tarkasteluihin. Tästä esimerkkinä käsitelen Paul Forresterin artikkelia ”The Desire Machine”. Artikkelissa kokemuskoneen pohjalta muodostetaan uusi ajatuskoe, halukone, jonka avulla haluteoriaa pyritään kehittämään. Paremman teorian tulisi hyväksyä ulkopuolinen standardi, halujen sopivuus, joka osittain määrittäisi sitä, kuinka paljon subjektin tyydyttynyt halu tuottaa hyvinvointia. Sopivat halut voidaan määritellä subjektin todellisista haluista riippumatta. Huomioin tutkielmassani myös sen, että kokemuskoneen uuden kyselytutkimuksiin nojaavan kritiikin myötä kokemuskoneen mallia voi mahdollisesti hyödyntää myös empiiristen kysymysten tarkasteluun.

Avainsanat

Ajatuskoe, haluteoria, hedonismi, hypoteesi, intuitio, kokemuskone, metodologia, Robert Nozick, subjektivismi, teoria, utilitarismi, vinouma

Sisällysluettelo

1. Johdanto	1
2. Hyvinvointiteorioita	2
2.1. Utilitarismi ja hedonismi	2
2.2. Haluteoria	3
3. Kokemuskone	4
3.1. Ajatuskoe teoksessa Anarchy, State, and Utopia	4
3.2. Argumentit utilitarismia ja hedonismia vastaan	5
4. Kritiikki: vinoumat häiritsevät kokemuskoneen tulkintaa	6
5. Ajatuskokeet filosofiassa	9
5.1. Argumentti tai intuitio	10
5.2. Ajatuskokeiden testaamisesta	10
6. Ajatuskokeet filosofisen teorianmuodostuksen apuna	11
6.1. Metodologisesti kiinnostavat tapaukset	11
6.2. Ajatuskoe vastaesimerkkinä teorialle	12
6.3. Ajatuskoe valmiina mallina	12
7. Miten luoda teoriaa kokemuskoneella?	13
7.1. Totuuteen mukautettu hedonismi	14
7.2. Kokemuskone mallina: Halukone	17
7.3. Kokemuskone: subjektiivisen ja objektiivisen välissä?	21
8. Yhteenveto: kokemuskoneen metodologisesta kiinnostavuudesta	22
Lähteet	24

1. Johdanto

Robert Nozickin (1938–2002) kokemuskone (engl. *The Experience Machine*) on tunnettu ajatuskoe filosofian alalla. Erilaiset kokemuskoneet lienevät vieläkin tunnetumpia populaarikulttuurissa esimerkiksi elokuvista, kuten *The Matrix* (1999). Kokemuskoneessa voit elää sellaista elämää, jota olet aina halunnut elää. Ainoa huonopuoli vain on, että kokemuksesi eivät ole todellisia. Miksi välittää siitä? Millä muulla voi olla väliä kuin miltä elämämme tuntuvat sisältäpäin?

Kokemuskone tunnetaan erityisesti sen kritiikistä utilitarismia ja hedonismia vastaan. Sen vaikutusvaltainen kritiikki on kuitenkin kyseenalaistettu uuden sukupolven tutkimuksen myötä (Lin 2016). Viisikymmentä vuotta sitten teoksessa *Anarchy, State, and Utopia* (1974) julkaistulla ajatuskokeella voi kuitenkin olla myös toisenlaisia käyttötarkoituksia.

Tässä kandidaatintutkielmassa tutkin sitä, miten kokemuskone voisi edesauttaa erilaisten filosofisten teorioiden ja hypoteesien muodostamista. Tarkastelun lähtökohdaksi olen valinnut Sara Kier Präemin ja Asbjørn Steglich-Petersenin artikkelin ”Philosophical thought experiments as heuristics for theory discovery” (2015). Präemin ja Steglich-Petersenin mukaan ajatuskokeita käsittelevässä kirjallisuudessa on jäänyt huomitta se, miten filosofiset ajatuskokeet ovat uusien teorioiden ja hypoteesien kannalta hyödyllisiä.

Ajatuskokeet voivat olla teorianmuodostuksessa hyödyllisiä siten, että ne tarjoavat vastaesimerkin, jonka ratkaiseminen vaatii jo olemassa olevan teorian korjaamista. Ajatuskokeet voivat myös tarjota valmiin mallin, jota voi hyödyntää muunlaisissa tarkasteluissa. Erityisesti ne voivat olla avuksi, kun muokkaamme alkuperäisiä ajatuskokeita.

Esittelen tutkielmassani Fred Feldmanin (2004) *totuuteen mukautettua asenteellista hedonismia* (engl. *Truth-Adjusted Intrinsic Attitudinal Hedonism*), joka on teoriana vastaus valheellisten nautintojen ongelmaan. Valheelliset nautinnot ovat yksi klassinen hedonismin kritiikki, jonka tunnetuin esimerkki kokemuskone on. Tässä hedonismin muotoilussa todelliset ja epätodelliset nautinnot erotetaan toisistaan. Nautinnon kohteen todellisuus on myös nautinnon itseisarvoon vaikuttava tekijä.

Esittelen myös tuoreemman julkaisun, Paul Forresterin artikkelin ”The Desire Machine” (2024), jossa kokemuskoneen mallin pohjalta muodostetaan uusi ajatuskoe, halukone. Halukoneen vastaesimerkin myötä Forrester ehdottaa, että haluteorian tulisi hyödyntää *sopivien halujen* -käsitettä. Kokemuskoneen uudempi kritiikki on nähtävissä Forresterin tavassa lähestyä omaa ajatuskoettaan.

Kyselytutkimuksiin nojaavassa kritiikissä kokemuskoneen erilaiset versiot tarjoavat lisäksi kiinnostavaa tietoa valinnoistamme ja arvostuksistamme.

2. Hyvinvointiteorioita

Tutkielman ymmärrettävyyden kannalta on hyvä käydä läpi siihen liittyviä hyvinvointiteorioita. Esittelen seuraavaksi utilitarismin, hedonismin ja halu- eli preferenssiteorioiden peruseriaatteita.

2.1. Utilitarismi ja hedonismi

Utilitarismi on tunnettu ja vaikutusvaltainen seuraustiikan muoto. Seurauseettisen näkemyksen mukaan asian tai teon moraalista arvoa tulee tarkastella niiden aiheuttamien seurausten perusteella. Mikäli teon X seuraukset ovat lopputulemaltaan parempia kuin teon Y, tulee valita X, ja hylätä Y. (Sinnot-Armstrong 2023.)

Klassisen utilitarismin mittapuuna tekojen hyvydelle on niiden seurausten aiheuttama kokonaishyvinvoinnin määrä. Teot, jotka maksivoivat hyvinvoinnin, ovat moraalisia tekoja. Hyvä ja hyvinvointi tarkoittavat käytännössä samaa. Relevantteja osapuolia vaikutusten arvioinnille ovat kaikki ne, joihin teon seuraukset vaikuttavat. Mikäli jonkin teon seurauksena hyvinvointi kasvaa hieman pienellä väestönosalla, mutta vähenee merkittävästi suuremmalla osalla, ei kokonaishyvinvointi todellisuudessa kasva, ja teosta on pidättäydyttävä. (Lagerspetz 2024, 170–172.)

Utilitaristien parissa hyvinvoinnin määrittelyyn on käytetty paljon aikaa. Tästä huolimatta yksimielisyyttä hyvinvoinnin (*utiliteetin* eli hyödyn) luonteesta ei ole syntynyt. Utilitarismia luonnehtii *arvomonismi*. Perustavin arvo on kaikkien tuntevien olentojen hyvä. Utilitaristisissa näkemyksissä hyvää tarkastellaan tekojen lopputulosten perusteella. Tarkastelun kohteena ei siis ole esimerkiksi se, tehtiinkö jokin teko halusta noudattaa jotakin moraalilakia. Lisäksi on olennaista, että lopputulokset eivät ole riippumattomia osatekijöistään. Näin ollen esimerkiksi valtion hyvinvointi on todellisuudessa sidottu sen kansalaisten hyvinvoinnin määrään. (Lagerspetz 2024.)

Klassiset utilitaristit, kuten Jeremy Bentham (1748–1832), John Stuart Mill (1806–1873) ja Henry Sidgwick (1838–1900) sitoutuivat hedonismiin eli hyvinvoinnin kokemusteoriaan. Itseisarvoista on nautinnon lisääminen ja tuskan välttäminen. Hyvinvointi, nautinto ja tuska liittyvät siis olennaisesti toisiinsa. (Sinnot-Armstrong 2023.) Benthamin mukaan kaikki hyvät sekä huonot kokemukset palautuvat nautinnon ja tuskan kokemuksiin. Mill sen sijaan kiinnittää huomion myös laadullisiin eroihin: ihmisellä on kyky kokea korkeamman tason nautintoja verrattuna eläimiin. Myöhemmät

1900-luvun utilitaristit painottavat enemmän onnellisuutta kuin nautintoa hyvinvoinnin mittana. (Lagerspetz 2024.)

Hedonismi on *subjektivistinen* näkemys hyvinvoinnin luonteesta. Nautinto, onnellisuus tai hyvinvointi on ennen kaikkea yksilön kokemus kyseisistä asioista. Hedonismi kuvaa sitä, mikä yksilön elämässä on arvokasta, eli miellyttävät kokemukset (Lagerspetz 2024, 176–177.) Utilitaristista ajattelua voi myös soveltaa muunlaisiin hyvinvointikäsitteisiin. Historiallisesti utilitarismi ja subjektivistinen hyvinvointikäsite kuitenkin liittyvät toisiinsa läheisesti. (Lagerspetz 2024, 190.)

2.2. Haluteoria

Halu- eli *preferenssiteorioissa* olennaista on se, että hyvinvointi määritetään lähinnä halujen tyydyttämiseksi. Haluteoria on myös subjektivistinen näkemys, jonka suosio on kasvanut hedonismin menetettyä suosiotaan. Hedonismista haluteoriat eroavat siten, että tyydyttyneen halun ei tarvitse johtaa mielihyvän kokemukseen. Halun tyydyttymistä ei ole välttämättä tarvitse kokea sisäisesti lainkaan. Halut suuntautuvat erilaisiin maailman asiantiloihin ja objekteihin, jotka voivat tyydyttää tai olla tyydyttämättä erinäisiä haluja. Hedonismissa vain subjektin sisäiset tilat ovat arvokkaita. Sisäisten tilojen arvo on riippumaton siitä, millaiset asiointitilat todellisuudessa vallitsevat. Haluteoriassa sen sijaan merkityksellistä on se, miten todellisuus suhteutuu subjektin haluihin. (Lagerspetz, 2024, 178–179.) Jos esimerkiksi haluaisin ystäväni lukevan filosofisia teoksia, hyvinvointini kasvaisi, mikäli he tosiasialisesti tekisivät näin. Hyvinvointini kasvaisi myös siinä tapauksessa, vaikka en koskaan saisi varmuutta siihen, onko haluni tyydyttynyt vai ei. Hedonistisessa katsantokannassa se, että ystäväni lukevat filosofiaa voi lisätä hyvinvointiani, jos koen sen tavalla, joka lisää mielihyvääni. Sinällään hyvinvointini ei ole kuitenkaan riippuvaista ystäväni todellisista lukumielityksistä. (ks. Lagerspetz 2024.)

Hyvinvoinnin ja halujen liittäminen toisiinsa ei ole ongelmaton. Olisi kummallista, jos hyvinvointia määrittäisi se, että hyvinvointi kasvaisi minkä tahansa halun tyydyttämisestä. Halut vaativat myös perustelua. Ne eivät perustele itse itseään, vaan pikemminkin haluamme asioita jonkin arvostamme tai hyvänä pitämämme asian tähden. Yksi mahdollinen ratkaisu ongelmaan on se, että hyvinvoinnissa on kyse *rationaalisten* halujen tyydyttämisestä. Mikäli ihmisellä olisi tarjolla riittävästi tietoa erilaisista vaihtoehdoista ja vaihtoehtojen seurauksista, häntä ei painostettaisi tai manipuloitaisi ja hänen päättelynsä olisi virheetöntä, voitaisiin hänen halunsa määritellä rationaaliseksi. Tässäkin piilee tietty ongelma: mitä täysin tiedoin varustettu rationaalinen olento haluaisi tai olisiko olennon

ja epätäydellisen ihmisen haluissa mitään samaa? (Lagerspetz 2024, 180–181.) Palaan rationaalisuuden ja halujen teemaan tutkielmani luvussa 7.2.

3. Kokemuskone

3.1. Ajatuskoe teoksessa *Anarchy, State, and Utopia*

On mielenkiintoista, miten Robert Nozickin kokemuskone elää ajassa. Lienee intuitiivista ajatella tekoälyaikakauden ihmisen kykenevän Nozickin ajatusleikkiin tavalla, joka poikkeaa 1970-luvun keskivaiheen maallikon käsityksistä: Erilaiset virtuaalitodellisuudet ovat kuluttajien saavutettavissa teknologisten laitteistojen kehityksen myötä. Elokvataiteeseen virtuaalitodellisuus on tehnyt läpimurron viimeistään Matrix-trilogian vaikutuksesta. Videopelitutkimukseen on tullut käsite immersio, joka tarkoittaa uppoutumista keinomaailmaan niin, että peliä ympäröivä todellisuus unohtuu. Teoksessa *Anarchy, State, and Utopia* (1974) ajatuskoe on seuraavanlainen:

Kysymykseen siitä, millä muulla on väliä kuin miltä *ihmisten* kokemukset tuntuvat heistä ”sisältöpäin”, liittyy huomattavia pulmia. Oletetaan, että olisi olemassa kokemuskone, joka voisi tarjota sinulle minkä tahansa kokemuksen, josta haaveilet. Alansa huiput neuropsykologit stimuloisivat koneen avulla aivojasi niin, että ajattelisit ja tuntisit kuin kirjoittaisit loistavaa romaania, ystäväystyisit jonkun kanssa tai lukisit mielenkiintoista kirjaa. Tämän ajan kuitenkin kelluisit tankissa elektrodit kytkettyinä aivoihisi. Menisitkö tällaiseen koneeseen ja ohjelmoisit elämäkokemuksesi ennakoita? Jos mietit, että jäisit paitsi toivottavista kokemuksista, voimme olettaa, että yritykset ovat perinpohjaisesti tutkineet monen muun elämää. Voit valita heidän laajasta valikoimastaan kokemuksia, sanotaan, kahdeksi vuodeksi eteenpäin. Kahden vuoden kuluttua sinulla on kymmenen minuuttia tai tuntia aikaa tankin ulkopuolella, jotta voisit valita kokemuksesi seuraavaksi kahdeksi vuodeksi. Tankissa ollessasi et tietenkään tiedä olevasi siellä; luulet kaiken todella tapahtuvan. Muut ihmiset voivat myös siirtyä kokemuskoneeseen, joten sinun ei tarvitse huolehtia heistä. (Jätä huomiotta ongelmat, kuten kuka huolehtii tankeista, jos kaikki siirtyvät kokemuskoneeseen.) Valitsisitko kokemuskoneen? *Millä muulla voi olla väliä kuin miltä elämämme tuntuvat sisältöpäin?* Sinun ei myöskään tulisi jättäytyä pois sen ahdingon vuoksi, joka ilmenee päätöksesi ja koneeseen siirtymisen välisenä aikana. Mitä muutama ahdingon hetki merkitsee verrattuna autuaaseen elämään (jos siis valitset siirtyä koneeseen), ja miksi tuntee ahdingkoa ollenkaan, jos päätöksesi on paras mahdollinen? (Nozick 2008, 42–43, oma suomennos)

Nozick uskoo, että ihmiset eivät siirtyisi kokemuskoneeseen. Kun kuvittelemme koneen, ymmärrämme myös, että emme käyttäisi sitä. Syiksi kieltäytyä hän mainitsee, että haluamme tehdä tiettyjä asioita sen sijaan, että vain tuntisimme tekevämme näitä. Elämä virtuaalitodellisuudessa on eräänlainen itsemurha: ikään kuin muulla ei olisi väliä kuin miten koemme itsemme omissa

tuntemuksissamme. Psykoaktiiviset huumausaineet ovat eräänlaisia reaalimaailman kokemuskoneita. Ei ole olemassa mitään syvempää todellisuutta, vaikka kokemus tällaisesta voidaan keinotekoisesti saavuttaa. Se, mikä on yhdelle syy tutkia syvempää todellisuutta, näyttäytyy toiselle kellumisena kokemuskoneen tankissa. (Nozick 2008, 43–44.)

3.2. Argumentit utilitarismia ja hedonismia vastaan

Kokemuskone on tulkittu argumenttina utilitarismia vastaan. Argumentti voidaan muotoilla esimerkiksi seuraavanlaisesti:

P1: Jos utilitarismi on totta, sinun tulisi siirtyä kokemuskoneeseen.

P2: Valitsisit kuitenkin oman elämäsi kokemuskoneen sijaan.

J: Utilitarismi ei ole totta. (Feldman 2011, 9.)

Useat ovat tulkinneet kokemuskoneen kritisoivan pikemminkin hedonismia. Argumentti hedonismia vastaan:

P1: Valitsisit oman elämäsi kokemuskoneen sijaan.

P2: Jos (P1), niin kokemuskone ei tuota enemmän hyvinvointia kuin oma elämäsi.

P3: Jos elämä kokemuskoneessa ei tuota enemmän hyvinvointia kuin oma elämäsi, niin hedonismi ei ole totta.

J: Hedonismi ei ole totta. (Feldman 2011, 11.)

Utilitarismin ja hedonismin lisäksi kokemuskoneen on katsottu haastavan hyvinvointiteoriat, jotka sitoutuvat mielentiloihin. Tämä tarkoittaa, että hyvinvoinnin määrä kasvaa tai laskee vain, mikäli ihminen tulee tietoiseksi siitä. (Kawall 1999, 382.) Jos jokin tapahtuma, kuten toisen ihmisen uhrautuminen puolestani, ei tule tietoisuuteni, ei tapahtumalla ole vaikutusta hyvinvointiini kumpaankaan suuntaan.

Tarvitsemme kuvittelukykyämme haaveidemme muodostamiseen sekä kokemuskoneen ymmärtämiseen. Tässä piileekin yksi kokemuskoneen hienouksista ja vaikeuksista. Ensin tulisi muodostaa jonkinlainen ajatus siitä, millaista elämä virtuaalitodellisuudesta olisi, ja tiedostaa ettei vaihtoehtoinen elämä toteudu. Mikäli vaihtoehtoisen elämän mahdottomuuden tiedostaa jo ennakolta,

voi olla vaikea ryhtyä koko ajatusprosessiin. Tosielämässä monen haaveen toteutuminen on vaihtokauppa tai sarja vaihtokauppoja. Näitä vaihtokauppoja on vaikea ennustaa etukäteen. Mick Jagger ei voi samaan aikaan olla rakastava perheenisä, joka kuljettaa lapsensa päivittäin jalkapalloharjoituksiin, ja maailmanluokan rock-tähti. Olennaista on, että emme yliajattele haaveidemme muodostavan jonkinlaisen loogisen ristiriidan, joka rikkoo todellisuuden ja kaataa koneen. Kokemuskoneessa meidän ei tarvitse keskittyä siihen, millainen todellinen potentiaalimme on suhteessa haaveisiimme.

Kokemuskoneen ymmärtäminen on kuitenkin osoittautunut haasteelliseksi. Kriitikot ovat argumentoineet, että intuitiomme ajatuskokeesta ovat vinoutuneet. (esim. Kolber 1994; Weijers 2014.). Kyselytutkimuksissa myös koehenkilöiden raportoimat syyt siitä, miksi he eivät valitsisi kone-elämää, on osittain annettu väärin perustein. Toisin sanoen ajatuskokeeseen on kuviteltu jotain siihen kuulumatonta, joka vaikuttaa valintaan. (Weijers 2014.)

4. Kritiikki: vinoumat häiritsevät kokemuskoneen tulkintaa

Filosofia ei ole empiirinen tiede, mutta kokemuskonetta on tutkittu myös kyselytutkimuksilla. Seuraavaksi esittelen sekä Felipe De Brigardin (2010) että Dan Weijersin (2014) tutkimuksia. Osion loppupuolella käsittelen myös lyhyesti Frank Hindriksin ja Igor Douven (2017) tutkimusta. Tutkimusten perusteella vaikuttaisi siltä, että ajatuskokeen ymmärtäminen voi osoittautua haasteelliseksi. Tämän lisäksi psykologiset taipumuksemme suosia nykyistä elämäämme ja omaa itseämme vaikuttavat, kun pohdimme kahden vaihtoehdon välillä.

Dan Weijers (2014) teetti tutkimuksen Wellingtonin Victoria-yliopiston opiskelijoille Nozickin kokemuskoneesta. Tutkimuksessa koehenkilöä pyydettiin vastaamaan, olisiko hän halukas siirtymään kokemuskoneeseen. Vastaukselle tuli antaa myös perustelut. Opiskelijat lukivat pääaineenaan joko filosofiaa tai kauppatieteitä. Kyselyyn vastasi 79 opiskelijaa, joista kolmetoista opiskelijaa vastasi myönteisesti. Mielenkiintoista on, että kielteisesti vastanneista 34 % (20/66) antoi perusteluksi valinnalleen seikkoja, jotka eivät olleet harmoniassa alkuperäisen ajatuskokeen skenaarion kanssa: Esimerkiksi se, että yllättävät kokemukset ovat parempia kuin valmiiksi ohjelmoidut. Lisäksi opiskelijat mainitsivat, että huonot kokemukset ovat välttämättömiä, jotta voisi arvostaa hyviä. Jotkut myös ajattelivat, että itse kokemuskone on pelottava tai luonnoton. (Weijers, 2014, 519–521.)

Weijersin kyselyn tulokset ovat mahdollinen esimerkki siitä, miten jotkut saattavat kuvitella kokemuskoneeseen liittyvän piirteitä, jotka eivät niihin kuulu. Weijers käyttää tästä termiä

kuvitteellinen vastarinta (engl. *imaginative resistance*), joka kuvaa enemmänkin vaikeutta lähestyä kuviteltua tilannetta sen itsensä kautta. (Weijers, 2014.)

Adam Kolber on argumentoinut, että intuitiomme kokemuskoneesta ovat erilaisia, mikäli Nozickin skenaario käännetään päinvastaiseksi. Kuvitellaan, että saat tietää eläneesi tähänastisen elämäsi koneen sisällä. Valinta on vaikea, sillä tässä tapauksessa yksilön tulisi luopua identiteetistä ja elämästä, joihin hän on jo tottunut. Mikäli jäämme virtuaalitodellisuuteen, on tilanne vastaava kuin se, että kieltäytyisimme Nozickin kuvittelemasta asetelmasta. Kolberin mukaan ajatuskoetta vaivaa niin sanottu *status quo -vinouma* (engl. *status quo bias*), eikä se siksi ole argumenttina pätevä. (Kolber 1994, 15.)

Yksi tapa tarkastella kyseistä vinoumaa on evolutiivinen. Evolutiivisesta näkökulmasta ihmiset ovat epävarmassa tilanteessa taipuvaisia tappiokammoon. (Weijers 2014, 522; Chen et al. 2006.) Kun emme tiedä, mitä kokemuskoneeseen siirtymisestä seuraa, haluamme asioiden pysyvän ennallaan. Suosimme sitä, että nykyinen elämämme ei muutu.

On toki luontevaa ajatella, että myös sillä on väliä, millainen elämämme on sen jälkeen, jos päätämme poistua koneesta. Felipe De Brigardin tutkimuksessa annettiin osallistujille kolme erilaista kuvausta siitä, mitä odottaa todellisuudelta kone-elämän jälkeen. Neutraalissa tilanteessa lisätietoja yksilön olemuksesta todellisuudessa ei annettu. Negatiivisessa tilanteessa koehenkilö sai tietää olevansa todellisuudessa vanki korkean turvallisuusluokan vankilassa Länsi-Virginiassa. Positiivisessa tilanteessa sen sijaan koehenkilö on Monacossa asuva taitelija, joka sattuu olemaan monimiljonääri. (De Brigard, 2010, 46–48.)

De Brigardin tutkimuksessa osallistujat jaettiin kolmeen ryhmään. Jokainen ryhmä vastasi yhteen yllä esitettyyn skenaarioon. Vastauksia kerättiin yhteensä 24 kappaletta yhtä kuviteltua tilannetta kohden. Negatiiviseen asetelmassa, jossa koehenkilö on todellisuudessa vanki, ainoastaan 13 % vastaajista oli halukas poistumaan virtuaalitodellisuudesta. Monacon taiteilijamiljonäärin ja kokemuskoneen välillä vastauksissa saavutettiin tasatilanne. Puolet vastaajista olivat halukkaita jättämään nykyisen elämänsä kokemuskoneessa, kun taas toinen puoli suosi jäämistä. Kun tilanne oli neutraali, vastaajista 54 % valitsi todellisuuden. (De Brigard, 2010.)

Weijers pyrki tutkimuksissaan vähentämään kuvitteellista vastarintaa ja status quo -vinoumaa uusilla ajatuskokeen muotoiluilla. Hän huomasi, että muokatut versiot myös tuottivat erilaisia tuloksia verrattuna Nozickin alkuperäiseen versioon. Esimerkiksi se, että ajatuskokeessa taattiin koneen varma toiminta, autonomian tunne virtuaalitodellisuuden sisällä, mielettömän hyvät kokemukset ja

muut kyselyyn osallistuneita huolettaneet seikat, innosti opiskelijoiden halukuutta siirtyä virtuaaliodellisuuteen 18 %. (Weijers 2014, 521–527.)

Jotta mahdollinen status quo -vinouma ajatuskoetta häiritsevänä tekijänä voitaisiin poistaa, Weijers pyysi opiskelijoita pohtimaan valintaa heille entuudestaan tuntemattoman ihmisen kanssa. Oma mielipidettä ei ollut tarkoitus kertoa tuntemattomalle, mutta valinta tuli tehdä siitä, pitäisikö vieraan ihmisen valita elämä kokemuskoneessa. Kun valinnassa ei ollut kyse omasta päätöksestä, jopa 52 % (43/82) koki, että kokemuskone parantaa elämänlaatua. Sen sijaan osallistujat olivat haluttomampia suosittelemaan keinoelämää ihmisille, jotka he kokivat läheiseksi, kuten serkulle (45 %) tai ystävälle (43 %). Weijersin mukaan eron tulosten välillä selittävät evolutiiviset tekijät: suosimme meille läheisiä ihmisiä tuntemattomien sijaan. (Weijers 2014.)

Suurin ero tutkimuksessa kahden eri version välillä saatiin, kun verrattiin Nozickin alkuperäistä versiota tilanteeseen, jossa valinta tehtiin tuntemattoman ihmisen puolesta. Lisätiedoksi annettiin se, että tuntematon oli viettänyt puolet elämästään kokemuskoneessa ja puolet todellisuudessa. Kokemukset todellisuudesta olivat pääsääntöisesti ikäviä, kun taas elämän parhaat kokemukset olivat simulaation tuotosta. Alkuperäisessä skenaariossa jo mainitut kolmetoista (13/79) opiskelijaa olivat myöntäväisiä siirtymään kone-elämään. Muokattu versio, jossa tuntemattoman todellista elämää luonnehtivat ikävät kokemukset, sai sen sijaan 42(/77) myönteistä vastausta. Ero oli siis tilastollisesti merkittävät 38 %. Huomioitavaa on, että lisätiedot siitä, millaista elämä todellisuudessa oli, ei tuntemattoman ihmisen kohdalla vaikuttanut juurikaan valintaan. Näin ollen De Brigardin hypoteesi erilaisten todellisuuskvausten vaikutuksesta valintaan liittyy lähinnä valintoihin, jotka liittyvät vain omaan itseemme. (Weijers 2014, 527–529.)

Weijers päättelee, että kokemuskone ei ole ratkaiseva argumentti hedonismia tai muita mielentiloihin sitoutuvia hyvinvointiteorioita vastaan tai ainakaan sitä ei tulisi tulkita sellaiseksi. Toki se tarjoaa jotain todistusaineistoa teorioita vastaan, mutta sen ongelmana on, että status quo -vinouma sekä muut epäoleelliset tekijät häiritsevät tulkintaa. Se, että tuntemattoman kohdalla ajatuskoe koetaan hyvänä vaihtoehtona, ei tietenkään tarkoita, että hedonismi olisi teoriana oikea. Sen sijaan on huomioitava, että tässä tilanteessa todellisuutta suosivat ovat vähemmistö. Kyseisten syitten vuoksi kokemuskone ei ole hedonismin ratkaiseva kritiikki. (Weijers 2014, 529–530.)

Weijersin mukaan on ongelmallista, että opiskelijat eivät ole tietoisia heidän arvioihinsa vaikuttavista taustatekijöistä. Teorioiden vastustamisessa itsessään ei ole vikaa, mutta epärealistisen ajatuskokeen olettaminen merkittäväksi kritiikiksi ilman, että arviointia häiritsevät tekijät tiedostetaan, ei ole

perusteltua. Lopuksi Weijers ehdottaa, että taustatekijöiden huomioimisen tulisi olla olennainen osa filosofian opetusta. (Weijers 2014.)

Hindriksin ja Douvenin tutkimustulokset olivat linjassa sekä De Brigardin että Weijersin tutkimusten kanssa. Tutkimukseen osallistuneiden määrä oli kyseisessä tutkimuksessa suurempi verrattuna aiempiin tutkimuksiin. Ensimmäiseen kyselyyn saatiin 224 vastausta ja toiseen 231. Hindriks ja Douven lisäsivät tutkimukseensa kaksi muuta mahdollista skenaariota: kokemuspillerin ja toimintakykypillerin. (Hindriks & Douven 2017.)

Kokemuspilleri muutti kaikki ihmisen loppuelämän kokemukset nautinnollisiksi. Toimintakykypilleri sen sijaan paransi ihmisen terveyttä, sekä kognitiivisia että sosiaalisia taitoja. Heidän motivaationsa lisätä tutkimukseen kaksi uutta koasetelmaan oli *autenttisuuden intuition* (engl. *The Authenticity intuition*) testaaminen: Ihmiset arvostavat elämää, joka on kosketuksessa todellisuuteen. Lisäksi ihmiset välittävät siitä, mitä he tekevät ja keitä he ovat. Tämä autenttisuus vaihtelee asteittain. Kyse ei ole siitä, että ihmiset joko arvostaisivat omaa itseään ja todellisuutta täysin tai lainkaan. (Hindriks & Douven 2017.)

Tutkimuksessa kokemuskone osoittautui kaikista vähiten houkuttelevammaksi vaihtoehdoksi. Kokemuspilleri innosti osallistujia kokemuskonetta enemmän. Toimintakykypilleri, joka oli eniten kosketuksessa todellisuuteen, osoittautui vaihtoehdoista suosituimmaksi. Tulokset tulkittiin autenttisuuden intuitiota tukevaksi. (Hindriks & Douven 2017.)

Kyselytutkimuksia on kritisoitu. Onhan totta, että kyselyihin vastanneet ihmiset eivät ole oikeassa valintatilanteessa. (Smith 2010, 38–39.). Käsittelen seuraavaksi sitä, miten filosofiset ajatuskokeet voidaan ymmärtää. Kaikkia ajatuskokeita ei voi toisintaa tosielämän kokeilla. Tämä ei välttämättä ole myöskään tarpeellista.

5. Ajatuskokeet filosofiassa

Ajatuskokeita muodostetaan filosofian lisäksi muilla tieteenaloilla, kuten matematiikassa, fysiikassa, historiassa ja biologiassa. Kuvittelukyky on keskeistä ajatuskokeille. Niillä on myös erilaisia käyttötarkoituksia. Ajatuskokeita voidaan käyttää esimerkiksi käsiteanalyysissa, hypoteesien muodostamisessa, teorian valitsemisessa ja opetuksessa. Kiistellyin ja ehkä keskeisin kysymys on, miten kuviteltuja tilanteita voidaan käyttää todellisuuden tutkimiseen. Tästä seuraa ajatuskokeiden yhteys muihin keskeisiin filosofisiin kysymyksiin: millainen on fiktion ja todellisuuden rajapinta,

mitä on kuvittelukyky tai intuitio, ja kuinka tärkeää on ymmärtää pikemminkin kuin selittää. (Brown & Fehige 2023.)

5.1. Argumentti tai intuitio

Yksi mahdollinen tapa tarkastella ajatuskokeita on, että ne ovat argumentteja. Ajatuskoe voidaan siis muotoilla argumentin muotoon (ks. 3.2.). Niitä voidaan käyttää argumentaation keinona, mutta niitä voidaan myös vastustaa samoin kuin argumentteja. Argumentti voidaan haastaa osoittamalla sen premissien, johtopäätöksen tai loogisen päättelyn ongelmallisuus. Ongelmaksi saattaa osoittautua se, jos filosofit tai muut tieteentekijät ovat erimielisiä siitä, minkä argumentin ajatuskoe oikeastaan esittää. (Brendel 2018, 281–282.)

Toinen tapa lähestyä ajatuskokeita on, että ne herättävät meissä *intuitioon* perustuvia arvioita ja johtopäätöksiä. Intuitioon turvautumista filosofiassa on arvioitu. Intuitioon turvautuminen ei ole yksinomaan filosofian alaa. Mikäli tarkastellaan sitä, mihin lopulta tietoiset johdatteluprosessit perustuvat, ei niiden takaa löydetä loputonta sarjaa toisia tietoisia johdatteluprosesseja, vaan lopulta päädyimme intuitioon: Fyysikko, joka tarkastaa kurinalaisesti laskelmansa ja tekee havaintoja siitä, että hänen päättelyprosessinsa pitää paikkansa, turvautuu myös intuitioon. Intuitioon turvautuminen ei tee hänen ajattelustaan irrationaalista, vaan pikemminkin jokin osa intuitiivisesta ajattelusta on myös osa rationaalista ajattelua. (Williamson, 2022, 62–68.)

5.2 Ajatuskokeiden testaamisesta

Filosofisia ajatuskokeita ei toisinneta todellisilla kokeilla, vaan ne jäävät usein kuvittelun tasolle. Osittain tämä johtuu siitä syystä, että filosofiset teoriat ovat usein jonkinlaisia yleistyksiä. Kuten muillakin aloilla, teoriaa voidaan testata tarkastelemalla sen seurauksia monissa mahdollisissa tilanteissa. Kaikkia mahdollisia tilanteita on kuitenkin mahdotonta kuvitella. Onkin haastavaa kuvitella skenaario, joka ennustaa teoriasta jotakin kiinnostavaa, eikä pelkästään jotakin triviaaliteettia. Mahdollisen tilanteen keksiminen ja sen avulla teorian seurausten selvittäminen on sinänsä jo ajatuskoe. Kun ajatuskoe esitetään vakavana kritiikkinä teorialle, on tärkeää, onko kuviteltu skenaario mahdollinen. Joidenkin ajatuskokeiden testaaminen todellisen maailman kokeilla on tietysti mahdotonta. (Williamson 2022, 59; 73–77.)

Vaikka filosofisilla ajatuskokeilla ja luonnontieteellisillä kokeilla on ilmeisiä eroja, niillä on myös yhteinen piirre: ne jättävät tilaa eri mielisyydelle. On tarpeellista varmistua siitä, ettei jonkin ennustuksen pieleen menemisessä ole vastuussa jokin muu kuin teoria itse. (Williamson 2022.)

Kriitikot ovat haastaneet kokemuskoneen argumentin vakuuttavuuden vetoamalla siihen, että ihmisten intuitiot muuttuvat, jos kuviteltua skenaariota muutetaan.

6. Ajatuskokeet filosofisen teorianmuodostuksen apuna

Artikkelissaan ”Philosophical thought experiments as heuristics for theory discovery” (2015) Sara Kier Präem ja Asbjørn Steglich-Petersen käsittelevät filosofisia ajatuskokeita jokseenkin toisenlaisesta näkökulmasta. Kirjallisuudessa ajatuskokeita on käsitelty näkökulmasta, jossa niiden rooli on pääasiassa joko teorian tai hypoteesin kumoaminen tai vahvistaminen. Huomiotta on jäänyt se, miten filosofialle tyypilliset ajatuskokeet mahdollisesti auttavat uusien teorioiden ja hypoteesien muodostamisessa. (Präem & Steglich-Petersen 2015, 1–2.)

Präem ja Steglich-Petersen ajattelevat, että syy sille, miksi näkökulmaa ei ole huomioitu, on hyvin ymmärrettävä. Lienee kiireellisempää saada vastauksia esimerkiksi sille, millaisen oikeutuksen filosofiset ajatuskokeet tarjoavat teorioiden puolesta tai niitä vastaan. Kokeiden usein intuitioon perustuvat arvostelmat on kyseenalaistettu epäluotettavina. Tästä on osittain kyse myös kokemuskoneen kritiikistä, jota olen tässä tutkielmassa tarkastellut. Toisaalta saattaa olla haastavaa tehdä hyvin systemaattinen esitys siitä, miten itse asiassa ajatuskokeet auttavat muodostamaan uusia teorioita. (Präem & Steglich-Petersen 2015.)

6.1. Metodologisesti kiinnostavat tapaukset

Mikä tahansa voi periaatteessa olla avuksi, kun muodostamme teorioita erinäisiin tarkoituksiin. Terveellinen ruokavalio, liikunta tai vaikkapa vieraiden ihmisten salakuuntelu ravintolan hämärässä nurkkauksesta voivat kaikki osaltaan stimuloida filosofista mieltä. Ajatus on hyvin intuitiivinen, mutta se ei varsinaisesti ole metodologisesti kiinnostava. Mikäli kriteerit ajatuskokeista teorioiden luojina ovat kovin väljät, on ilmiselvää, että ne edesauttavat uusien teorioiden luomisprosessia. (Präem & Steglich-Petersen 2015, 4). Onhan totta, että filosofista ajatusprosessia inspiroivat usein ajatukset, jotka ovat peräisin kymmenien sukupolvien takaa.

Jotta filosofiset ajatuskokeet olisivat metodologisesti kiinnostavia, tulisi niiden auttaa teorian muodostamisessa tavalla, joka ei ole liioin yllättävä. Tutkimustyön tulosten tulisi olla jossain määrin ennustettavissa tai ainakin jokin säännöllinen elementti on vaadittua. Se, että jokin ajatuskoe auttaa hylkäämään vain jonkin virheellisen teorian ja täten raivaa tilaa uusille tutkimuksille, ei ole metodologisesti kiinnostavaa. (Präem & Steglich-Petersen 2015.)

Präem ja Steglich-Petersenin mukaan kiinnostavien tapausten kriteeristöön ei kuitenkaan kuulu se, että tietty ajatuskoe on essentiaalinen sen suhteen, että jokin uusi teoria muodostuisi. On hyvin mahdollista, että teoria keksittäisiin jollain toisella tapaa. Toisaalta ajatuskokeen alkuperäisen funktion ei tarvitse olla nimenomaisesti uuden teorian esittäminen, vaan se voi olla esimerkiksi tietyn teorian vastustaminen. Ajatuskokeiden ei myöskään tarvitse olla mikään täydellinen metodi, kunhan niistä on apua jonkinasteisen säännöllisyyden puitteissa. (Präem & Steglich-Petersen 2015, 4–6.)

6.2. Ajatuskoe vastaesimerkinä teorialle

Ajatuskokeet voivat auttaa teorian muodostamisessa vastaesimerkkien myötävaikutuksesta. Tarkemmin vastaesimerkit mahdollistavat jo olemassa olevien teorioiden parantamisen. Ne voivat osoittaa teorian heikot kohdat. Näin pohdinnan voi suunnata erilaisiin räätälöityihin ratkaisuihin, jotka edistävät teorian uskottavuutta. Präem ja Steglich-Petersen nimittävät näitä ratkaisuja *ad hoc -korjauksiksi* (engl. *Ad hoc fixes*). (Präem & Steglich-Petersen 2015, 10–11.)

Vastaesimerkit voivat puuttua jonkin teorian tai sen osa-alueen riittävyteen (engl. *sufficiency*) tai välttämättömyyteen (engl. *necessity*). Esimerkiksi kuuluisat Gettier-tapaukset osoittivat sen, että klassisen tiedon määritelmän kolme ehtoa eivät ole riittäviä. Se, että tieto on (1) tosi (2) perusteltu (3) uskomus, ei sulje pois episteemisen sattuman mahdollisuutta. Tämä motivoi parantamaan kyseessä olevaa teoriaa tarjoamalla konkreettisen esimerkin. Paremman analyysin pitäisi siis huomioida tämä tekijä ja poissulkea sen mahdollisuus. (Präem & Steglich-Petersen 2015, 10–13.)

Ei ole täysin selvää, miten kaikki riittävyteen tai välttämättömyyteen liittyvät vastaesimerkit voivat auttaa teorian kehittämisessä. Haasteelliseksi tämä voi muodostua, mikäli ajatuskoe haastaa jonkin teorian fundamentaalisen oletuksen, eikä pelkästään jotakin teorian osa-aluetta. Kuuluisa ajatuskoe neurotieteilijä Marysta lienee hyvä esimerkki tällaisesta. Elämänsä mustavalkoisessa huoneessa asunut neurotieteilijä on lukenut kaiken mitä väreistä on ikinä kirjoitettu. Kun hän lopulta poistuu huoneesta ja näkee maailman värit ensikerran, oppiiko Mary jotakin uutta väreistä? Mikäli intuitio siitä, että Mary todella oppii jotain uutta, pitää paikkansa, ei ole täysin selvää, miten fysikalistista teoriaa mielestä voisi parantaa vastaesimerkin pohjalta. (Präem & Steglich-Petersen 2015.)

6.3. Ajatuskoe valmiina mallina

Ajatuskokeet voivat olla hyödyllisiä teorioiden kannalta myös muuten kuin tarjoamalla vastaesimerkin, johon keskittyä. Ajatuskokeen alkuperäinen funktio voi hyvin olla vain vahvistaa käsitystä, josta kirjoittaja on vakuuttunut. On siis mahdollista, että ajatuskoe seuraa vasta kun sitä edeltävä pohdinta on jo valmis, eikä ajatuskoe itsessään ole teorian syy. Vaikka teoria edeltäisikin

ajatuskoetta, voi ajatuskokeen tarjoama asetelma olla malli (engl. *template*) muille kehittää erinäisiä teorioita. Tässä mielessä ajatuskokeet voivat olla avuksi filosofiselle teorialle erityisesti, kun teemme erilaisia muokkauksia jo olemassa oleviin ajatuskokeisiin. Ei ole myöskään välttämätöntä, että muokatun version pitäisi käsitellä juuri samaa filosofian osa-aluetta tai teoriaa kuin alkuperäisen. (Präem & Steglich-Petersen 2015, 13–14.)

Präem ja Steglich-Petersen antavat esimerkiksi Putnamin kaksoismaa-ajatuskokeen, joka alun perin on tarkoitettu semanttisen eksternalismin tueksi. Se on kuitenkin tarjonnut mahdollisuuden myös muunlaisille filosofisille tarkasteluille, kuten: riippuvatko *mielementilat* ulkoisista tekijöistä (McGinn 1977) tai ovatko *moraaliset ominaisuudet* luonnollisia luokkia? (Horgan & Timmons 1992).

7. Miten luoda teoriaa kokemuskoneella?

Tässä kohtaa tutkielmaani tarkasteluni keskittyy siihen, miten kokemuskone ajatuskokeena voisi auttaa filosofisten hypoteesien ja teorioiden muodostamisessa. Hyödynnän analyysissäni edellä esittelemääni Präemin ja Steglich-Petersenin artikkelia. Aloitan keskittymällä heidän ajatukseensa siitä, että ajatuskokeiden tarjoamat vastaesimerkit voivat osoittaa teorioiden heikot kohdat, jotka vaativat lisätarkastelua tai korjausta. Monet ovat muokanneet hedonistista teoriaa kokemuskoneen pohjalta (Berkovski 2012, 186). Jotkut ovat argumentoineet, että kokemuskone ei tosiasiallisesti ole hedonistin valinta (esim. Bramble 2016, 101–104; Hewitt 2009, 384). Teorian muokkaamisesta hyvä esimerkki on Fred Feldmanin *totuuteen mukautettu asenteellinen hedonismi* (engl. *Truth-Adjusted Intrinsic Attitudinal Hedonism*), jota tarkastelen yksityiskohtaisemmin. (Feldman 2004.)

Tätä tarkastelua seuraa ajatuskokeen käyttö eräänlaisena mallina filosofiselle teorialle. Kokemuskone on ikään kuin sapluuna, jota on mahdollista laajentaa muunlaisille filosofisille tarkasteluille. Esimerkkinä kokemuskoneen käytöstä mallina esittelen Paul Forresterin (2024) *halukonetta*, jossa kokemuskoneen pohjalta luodaan uusi ajatuskoe. Halukoneen tarkoitus on kehittää haluteoriaa siten, että teoria vastaisi paremmin halukoneen osoittamaan kritiikkiin. (Forrester 2024.) Tarjoan myös lyhyesti muita esimerkkejä kirjallisuudesta, joissa kokemuskoneen tarjoamaa mallia hyödynnetään.

Teorioiden esittelyn jälkeen vertaan näkemyksiä keskenään. Kumpikaan teoria ei hylkää perusteoriaansa, vaan kehittää niitä ajatuskokeen avulla. Ratkaisut vaativat sitä, että nautintoa tai haluja ei voida arvioida pelkästään subjektiivisesta näkökulmasta. Esitän myös oman näkemykseni siitä, millaisia kysymyksiä tarkastelemani kirjallisuus herättää.

7.1. Totuuteen mukautettu hedonismi

Kokemuskone osoittaa, että täysin keinotekoisesti luodut kokemukset voivat johtaa elämään, joka on kaikkein nautinnollisin. Koska hedonismissa nautinto on itseisarvoista, voi olla oikein tavoitella miellyttäviä kokemuksia ilman, että kokemus perustuisi todellisiin tapahtumiin. Kysymys kuuluukin, millainen arvo todellisuudella hedonistille voi olla? Hedonismi on monistinen näkemys, joten todellisuudella voi korkeintaan olla välinearvoa nautinnon tavoittelemisessa.

Yksi mielenkiintoisimmista yrityksistä vastata kysymykseen on Fred Feldmanin muotoilema *totuuteen mukautettu asenteellinen hedonismi* (engl. *Truth-Adjusted Intrinsic Attitudinal Hedonism*). Teorian perustana on *asenteellinen hedonismi* (*Intrinsic Attitudinal Hedonism*). Tässä hedonismin muotoilussa on olennaista tehdä ero puhtaiden aistinautintojen ja asenteellisten nautintojen välille. Ihminen voi kokea muutosta asenteellisissa nautinnoissaan ilman, että se aiheuttaisi muutosta aistinautinnoissa. Ajatuskoe selventää asiaa. (Feldman 2004, 4.1.)

Kuvitellaan, että neurotieteilijä on pitkään yrittänyt kehitellä pilleriä, joka poistaa kyvyn kokea mitään nautintoa tai kipua aistien kautta. Kun hän viimein testaa itsellään lupaavalta vaikuttavaa pilleriä, tieteilijä onnistuu tavoitteessaan. Hajut, tunteet ja maut muuttuvat yhtäkkiä täysin neutraaleiksi. Hän riemastuu läpimurrostaan. Toisin sanoen hänen asenteellinen nautintonsa kasvaa sen tosiasian seurauksena, että hän onnistui poistamaan itseltään kyvyn kokea aistinautintoja. Jonkin ajan kuluttua vaikuttaa siltä, että pillerin vaikutus on odotettua pidempi. Tämän seurauksena herää huoli, että hän on tehnyt jotakin peruuttamatonta aisteilleen. Asenne uutta tilannetta kohtaan alkaa hiljalleen muuttua: aiempaan verrattuna hänen tieteellinen läpimurtonsa ei tunnu läheskään siltä, miltä se alun perin tuntui. Mikäli tilanne ei muutu parempaan, mielipaha valtaa hänen ajatuksensa. (Feldman 2004, 4.2.)

Asenteellisessa hedonismissa arvoa määräävät ajanjaksot, joissa asenteemme erinäisiin asiointiloihin tuottavat nautintoa tai tuskaa. Tällaiset jaksot arvioidaan niiden voimakkuuden ja keston perusteella. Voimakkuuden ja keston perusteella määritellään nautinnon tai tuskan kokonaismäärä. Neurotieteilijän nautinto alkaa, kun hän onnistuu pillerin kehittämisessä ja päättyy kun mielipahasta tulee hallitseva tunne. Jakson voimakkuus taas perustuu nautinnon keskiarvolliseen määrään kyseisen jakson aikana. Feldman myöntää, että tällaisten jaksoiden määrittely voi olla vaikeaa jopa niiden kokijalle itselleen. Käytännön määrittelyvaikeuksista ei kuitenkaan seuraa se, että teoria olisi vailla minkäänlaista totuutta: milloin aistinautinto alkaa ja päättyy, on yhtä vaikea kysymys. Joka tapauksessa ne tulevat ja menevät sekä vaihtelevat voimakkuuksissaan. (Feldman 2004, 4.2.–4.3.)

Asenteellinen hedonismi voidaan muotoilla seuraavasti:

- (i) Jokainen jakso asenteellista nautintoa on itseisarvoisesti hyvää; jokainen jakso asenteellista tuskaa on itseisarvoisesti pahaa.
- (ii) Nautinnollisen jakson itseisarvoa määrittää sen sisältämän asenteellisen nautinnon kokonaismäärä; tuskaisen jakson itseisarvoa määrittää vastaavasti asenteellisen tuskan kokonaismäärä.
- (iii) Elämän itseisarvo määräytyy nautinnollisten ja tuskaisten jaksojen itseisarvojen (eli jaksojen sisältämän kokonaisnautinnon/tuskan) myötä niin, että yksi elämä on toista arvokkaampaa, jos ja vain jos yhden elämän asenteellisen nautinnon nettomäärä on suurempi kuin toisen. (Feldman 2004, 4.3.)

Asenteellinen hedonismi ei itsessään vielä tosin vastaa kokemuskoneen sille asettamaan haasteeseen. Virtuaalitodellisuudessa asenteitamme erinäisiin asiointiloihin määrittävät elektrodeja pitkin syötetyt sähköimpulssit. Sen sijaan Feldmanin on tarkoitus osoittaa, että asenteellinen hedonismi kykenee vastaamaan klassisiin hedonismin kritiikkeihin sen erinäisillä johdannaisilla. Valheelliset nautinnot ovat yksi klassinen kritiikki, ja kokemuskone on sen tunnetuin esimerkki. Asenteellinen hedonismi on mukautuva teoria, jonka johdannainen *totuuteen mukautettu asenteellinen hedonismi* tarjoaa ratkaisun niille, jotka karttavat hedonistista hyvinvointiteoriaa vedoten valheellisten nautintojen mahdollisuuteen. (Feldman 2004, 4.1, 5.1.)

Totuuteen mukautetussa hedonismissa tuskan välttäminen on itseisarvoista riippumatta siitä, ovatko sen aiheuttajat tosia tai epätosia (Feldman 2004, 5.1). Näkemys ei siis tee eroa siinä, aiheutuuko tuska tosielämän tai virtuaalitodellisuuden tapahtumista. Tuskaa on joka tapauksessa vältettävä. Feldman kysyy, jos empaattinen henkilö uskoo, että Somaliassa on nälänhätä, onko sen aiheuttama mielihäiriö erilaista riippumatta siitä, mikä tilanne Somaliassa todellisuudessa on?

Näkemyksen mukaan kivun ja nautinnon välillä vallitsee kuitenkin asymmetria. Siinä missä kipua aiheuttavien jaksojen itseisarvoa määrittää jakson kesto ja voimakkuus, nautinnollisten jaksojen itseisarvollinen asema määritellään lisäämällä sen kestoon ja voimakkuuteen kolmas tekijä, totuus. Nautinnon lähteet, jotka aiheutuvat jostakin todellisesta, jakson kokonaisnautinto kerrotaan yhdellä. Vastaavasti valheellisista lähteistä saatu kokonaisnautinto kerrotaan 0,1:llä.

Totuuteen mukautuva hedonismi muotoillaan seuraavasti:

- (i) Jokainen jakso sisäistä asenteellista nautintoa on itseisarvoisesti hyvää; jokainen jakso sisäistä asenteellista tuskaa on itseisarvoisesti pahaa.

- (ii) Nautinnollisen jakson itseisarvoa määrittää sen sisältämän asenteellisen nautinnon kokonaismäärä, joka on ekvivalentti totuuteen mukautettuun määrään nautintoa jakson sisällä; tuskaisen jakson itseisarvoa määrittää vain asenteellisen tuskan kokonaismäärä.
- (iii) Elämän itseisarvo määräytyy nautinnollisten ja tuskaisten jaksojen itseisarvojen (eli jaksojen sisältämän kokonaisnautinnon/tuskan) myötä niin, että yksi elämä on toista arvokkaampaa, jos ja vain jos yhden elämän asenteellisen nautinnon nettomäärä on suurempi kuin toisen. (Feldman 2004, 5.1.)

Teorian seurauksena siis todelliset nautinnon lähteet ovat kymmenkertaisesti parempia kuin valheelliset vastaavat, mutta nautinto on yhä ainoa asia, joka lisää itseisarvoa. Onko kyse yhä hedonistisesta teoriasta? Feldman vastaa kysymykseen myöntävästi: totuuden funktio teoriassa on samankaltainen kuin nautinnon voimakkuus ja sen kesto. Muutokset sekä nautinnon voimakkuudessa ja kestossa vaikuttavat siihen, kuinka itseisarvoisia ne ovat. Se, minkälaiset objektit tai asiointilat ovat asenteellisen nautinnon lähteenä, vaikuttavat myös siihen, miten tavoiteltava tai itseisarvoinen jokin nautinto on. Epätodellisiin asioihin perustuvat nautinnot ovat yhä arvokkaita, mutta ne arvioidaan eri tavalla. (Feldman 2004.)

Perinteisesti hedonistinen oletus on, että kahden ihmisen mentaalisten tilojen ollessa samat, niiden arvo yhtä suuri (Feldman 2004, 8.4). Toisin sanoen, mikäli virtuaalitodellisuudessa ja todellisessa maailmassa kahden eri ihmisen mentaaliset tilat ovat yhdenmukaiset, heidän elämänsä on yhtä arvokasta. Nautinnoilla on kuitenkin eri kohde. Virtuaalitodellisuudessa nautinnon kohteena on ihminen itse, kun taas todellisessa maailmassa nautinto suuntautuu sen olemassa oleviin objekteihin. Keinotekoisen maailman tarkastelija voi kokea kyllä suurta nautintoa, mutta se johtuu hänen sisäisistä harhakuvitelmistaan. Todellisen maailman harhakuvitelmat ovat niin ikään itseisarvoa vähentäviä tekijöitä. Sille ei siis ole väliä, kuvitteleeko ihminen olevansa Mick Jaggerin kaltainen karismaattinen henkilö simulaation aiheuttamana, vai onko ajatuksesta seuraava mielihyvä puhdasta narsismia vailla totuutta. Kärsimystä on kuitenkin vältettävä riippumatta siitä, mikä sitä aiheuttaa. (ks. Feldman 2004, 4.1, 8.4.)

Totuuteen mukautettua hedonismia on syytetty siitä, että se on *ad hoc*, eli puhtaasti hedonistisen teorian pelastamiseksi laadittu teoria (Olsaretti 2006, 413–414). Toisaalta juuri yksi tapa, millä Präem ja Steglich-Petersen ajattelevat ajatuskokeiden luovan filosofista teoriaa, ovat heidän mainitsemansa Ad hoc -korjaukset. Feldmanin teoriaan liittyy monia käytännön vaikeuksia: esimerkiksi jaksojen pituuksien määrittely, keskiarvojen mittaaminen ja se, millainen painoarvo tosiasioille annetaan. Niitä voisi tietysti painottaa enemmän tai vähemmän. Ajattelen kuitenkin, että tärkein kysymys liittyy juuri siihen, millainen arvo näillä karkeasti kahdella erilaisella nautinnon lähteellä on. Monimutkaisessa

maailmassa hedonismi tarjoaa kuitenkin verrattain selkeitä arviointiperusteita asioiden eettiselle ja arvottavalle tarkastelulle. Teoria voi toimia eräänlaisena kompromissiratkaisuna niille, jotka kokevat hedonismin olevan muutoin intuitiivinen ja kannatettava teoria, mutta pitävät epäintuitiivisena ajatusta siitä, että todelliset ja epätodelliset nautinnot olisivat yhdenvertaisia. Ovatko esimerkiksi tekoälyyn perustuvat virtuaalittöystävät ja oikeat seurustelukumppanit yhtä arvokkaita, jos ne tuottavat saman määrän nautintoa? (esim. BBC, 2024.)

7.2. Kokemuskone mallina: Halukone

Toinen tapa, miten ajatuskokeet voivat auttaa filosofisten teorioiden muodostamisessa, on ajatuskokeen käyttäminen eräänlaisena mallina. Erityisesti niistä voi olla hyötyä silloin, kun teemme muokkauksia jo olemassa oleviin ajatuskokeisiin. (Präem & Steglich-Petersen 2015, 13–14.) Kokemuskone ja sen erilaisia versioita on hyödynnetty esimerkiksi kyselytutkimuksissa, joita tutkielmassa on käsitelty. Esimerkiksi Weijers (2014) on tutkimuksensa pohjalta argumentoinut, että ajatuskokeiden arviointia häiritsevät taustatekijät tulisi huomioida filosofian opetuksessa. Hindriksin ja Douvenin (2017) tutkimuksesta erilaisia versioita tarkasteltiin autenttisuuden intuition näkökulmasta.

Kyseiset tutkimukset ovat kokeellisen filosofian tai moraalipsykologian tutkimuksia. Tämän lisäksi kokemuskonetta on hyödynnetty perinteisen filosofian keinoin: Ovatko esimerkiksi suoratoistopalvelu Netflixin yhä enemmän yksilön mieltymyksiä mukailevat ohjelmasuositukset eräänlaisia reaali maailman kokemuskoneita? Tärkeää on, että suosituksista vastaava algoritmi toimii siten, että käyttäjä viihtyy ja jatkaa palvelun käyttöä. Personalisoidun käyttäjäkokemuksen kääntöpuolena on tosin se, että käyttäjän kannalta epämieluiset aiheet, tai millaista sisältöä toisenlaiset ihmiset katsovat, jäävät huomiotta. (Brincker 2021.) Tai miten tulisi suhtautua skitsofreniaa sairastavan harhoihin, jos ne eivät tee hänestä sen enempää onnetonta kuin onnellistakaan? Millaisin perustein muiden tulisi poistaa hänet vastentahtoisesti tällaisesta rikkinäisestä kokemuskoneesta? (Birks 2011.)

Artikkelissaan ”The Desire Machine” Paul Forrester muodostaa ajatuskokeensa *halukone* hyödyntäen kokemuskonetta. Vastoin alkuperäistä ajatuskoetta halukoneeseen ei liity vaihtoehtoista virtuaalitodellisuutta. Sen sijaan halukone on yksinkertainen neurokirurginen operaatio, joka saa sinut haluamaan niitä tapahtumia ja asiointiloja, jotka todellisuudessa vallitsevat. Ennen operaatiota tyydyttymättä jääneet halut ovat heikentäneet elämänlaatuasi, mutta poistuttuasi leikkaussalista kaikki turhautumisen tunne on kadonnut. Koska nyt suhtaudut myötämielisesti kaikkeen ympärilläsi tapahtuvaan, sinulla ei myöskään ole motivaatiota muuttaa mitään. (Forrester 2024, 1.)

Kokemuskone ei ole ongelma haluteorialle, sillä vaihtoehtoiseen todellisuuteen siirtyminen ei välttämättä ole yksilön halujen mukaista. Halujen tyydytys on hyvinvointia määrittävä tekijä, eikä se, miten paljon nautintoa valinnasta seuraa. (Forrester 2024, 2.)

Forrester tarkastelee halukonetta *prudentialisesta* ja *objektiivisesta* näkökulmasta, eli olisiko yksilön järkevää valita halukone. Kyse ei siis ole siitä, mitä toimija itse raportoisi sen suhteen, onko halukone hyväksi hänelle. Forrester esittelee artikkelissaan kaksi mahdollista haluteoreetikon vastausta halukoneen ongelmaan, halujen muutoksen (engl. *Change of Desire*) ja konatiivisen koherenssin (engl. *Conative Coherence*), mutta hylkää ne ongelmallisina. Tämän jälkeen hän ehdottaa muokkausta haluteorialle, joka vastaisi paremmin hänen omaan ajatuskokeeseensa. Ratkaisuna ovat *sopivat halut* (engl. *fitting desires*): tyydyttynyt halu tuottaa sitä enemmän hyvinvointia, mitä sopivampi se on. (Forrester 2024.)

On olemassa kaksi tapaa tyydyttää halumme. Yksi tapa on muuttaa maailmaa niin, että jo olemassa olevat halumme tyydyttyvät. Toinen vaihtoehto on muokata halujemme siten, että ne ovat yhdenmukaisia sen suhteen, millainen maailma jo nyt on. Halukone on siis ääriesimerkki tästä toisesta tavasta tyydyttää haluja. (Forrester 2024, 2–3.)

Operaation seurauksena halumme muuttuvat, sillä harvoin olemme tyytyväisiä kaikkeen mitä ympärillämme tapahtuu. Toisaalta taas on empiirinen tosiasia, että ihmisen halut muuttuvat hänen elämänsä aikana muutenkin. Mutta ovatko halut sellaisia, että ne vain muuttuvat ajan saatossa, vai voiko yksilöllä olla rationaalisia perusteita muuttaa halujaan? Ainakin voi olla rationaalista muuttaa uskomuksiaan silloin, kun saa uutta tietoa. (Forrester 2024, 1–4.)

Halujen muutoksen rationaalisuutta voidaan tarkastella Forresterin mukaan sekä muutosta ennen että sen jälkeen. Jos haluamme vaihtaa nykyisiä halujamme, tulisi muutoksen rationaalisuutta tarkastella hyödyn, eli *utiliteetin* näkökulmasta. Nykyisten halujen utiliteetti voidaan määrittää vertaamalla haluja siihen, miten todennäköisesti henkilö uskoo halujen tyydyttyvän. Odotettavissa oleva utiliteetti vaikuttaisi olevan matala tilanteissa, joissa henkilöllä on erilaisia haluja, jotka liittyvät samaan asiaan tai aktiviteettiin: esimerkiksi jos haluan lopettaa kahvinjuonnin aamuisin, mutta tästä huolimatta keitän itselleni aamukahvit. Toisaalta taas halun menettäminen vähentää todennäköisyyttä sille, että jokin asia, esimerkiksi kahvinjuonnin lopettaminen, tapahtuisi. (Forrester 2024, 4–5.)

Koska halun menettäminen vaikuttaa siihen, miten todennäköisesti jokin tapahtuu, halun menettämisestä koituu haittaa. Jos taas pidämme nykyiset halumme, jatkamme yhä asioiden tavoittelua, jotta halumme tyydyttyisivät. Tästä syystä melkein jokainen luonnollinenkin muutos

haluisamme voi olla irrationaalista: ehkäpä meidän ei olisi pitänyt koskaan luopua lapsuutemme haaveista. (Forrester 2024.)

Halujen muutoksen rationaalisuutta voidaan arvioida myös muutoksen jälkeen. Jos ajatellaan tilannetta, jossa henkilö voi valita vain yhden kahdesta halusta, ja molempien halujen utiliteetti on yhtä suuri, mikä ratkaisisi valinnan? Mikäli yksi haluista on realistisempi, helpommin toteutettavissa, olisi järkevää muodostaa kyseinen halu. Esimerkiksi, jos ihmisellä on halu matkustaa Marsiin tai opetella pelaamaan bridgeä, rationaalista on valita kahdesta halusta korttipeli punaisen planeetan sijasta. Tosin jos halut, jotka tyydyttyvät helpommin, ovat parempia valintoja, sitten jo tyydyttyneet halut ovat parhaita haluja. Halukone saa ihmisen haluamaan kaikkea, mitä hänen ympärillensä tapahtuu. Olisi siis rationaalista valita se, mikäli halujen muutosta arvioidaan vasta muutoksen jälkeen. (Forrester 2024, 5.)

Toinen mahdollinen vastaus, jonka haluteorian kannattajat voisivat antaa vastaukseksi halukoneen ongelmaan, on se, että halujen muutos on rationaalista, jos se edistää halujen koherenttiutta. *Konatiivinen reflektiivinen tasapaino* (engl. *conative reflective equilibrium*) on sisäinen tila, jossa ihmisen halut ovat mahdollisimman koherentteja keskenään ja eivät sisällä ristiriitaisia haluja. Tähän tilaan ihmisen tulisi pyrkiä halujensa suhteen. (Forrester 2024, 6–8.)

Sille, mikä tekee halujen verkostosta koherentteja, voidaan asettaa erilaisia ehtoja. Ehdot rajoittuvat siihen, mikä tekee henkilön subjektiivista haluista koherentteja. Sinänsä halujen sisältöön ei oteta kantaa. Esimerkiksi koherentissa halujen järjestelmässä ihmisen ei tulisi haluta asioita, joiden tietää olevan mahdottomia. Kriteeriksi voitaisiin muun muassa asettaa se, että uusien halujen tulisi olla entisten halujen koherentteja jatkeita (engl. *Coherent extensions*). (Forrester 2024, 7.)

Halukone sovittaisi halumme koherenteiksi keskenään, mutta muutos vanhan ja uuden tilan välillä olisi liian suuri. Koherenttiuden vaatimus ratkaisisi kyllä halukoneen ongelman, mutta koherenttiuteen liittyy muita ongelmia: Esimerkiksi henkilön halut ovat koherentteja myös siinä tilanteessa, että henkilöllä ei ole haluja lainkaan. On intuitiivista kuitenkin ajatella, että haluja tulisi olla. Vastaavasti Stalinin halu lähettää vastustajansa gulagiin on linjassa hänen aiempien halujensa kanssa. (Forrester 2024, 8–9.)

Halujen merkittävään muutokseen vetoaminen johtaisi halukoneen hylkäämiseen. Tämä tosin vastaisi vain kysymykseen, miksi halukonetta ei tulisi valita. Tästä huolimatta elämä halukoneessa haluteorian mukaan tuottaisi enemmän hyvinvointia. (Forrester 2024, 9–10.) Eden Lin on argumentoinut, että kokemuskone -tyyppisiä ajatuskokeita tulisi nimenomaisesti tarkastella hyvinvoinnin näkökulmasta: ei siis siten, että siirtyisikö koehenkilö kokemuskoneeseen tai pitäisikö

hänen niin tehdä (Lin 2016). Rationaalisten vaatimusten liittäminen hyvinvointiin koherentissa lähestymistavassa on ongelmallista: mikä subjektiivisessa näkökulmasta tarkasteltuna olisi järkevä syy jättää jotain tekemättä kuin että sillä olisi vaikutuksia ihmisen hyvinvointiin? (Forrester 2024.)

Koherentissa lähestymistavassa on tyypillistä erottaa halut muista mentaalisista tiloista, kuten uskomuksista. Halujen oikeellisuus (engl. Correctness) määritetään subjektista käsin. Ei siis ole subjektin ulkopuolista standardia, jonka kautta määrittää niiden oikeellisuus. Halun muuttaminen on rationaalista, jos ja vain jos se edistää konatiivista reflektiivistä tasapainoa. (Forrester 2024.)

Halukoneen ratkaisu vaatii muokkausta siihen, millaiseksi halut käsitetään. Halut tulisi käsittää aivan kuten muutkin intentionaaliset mentaaliset tilat. Uskomuksilla, peloilla, kateudella ynnä muilla mentaalisilla tiloilla on olemassa oikeellisuuden standardi (engl. Standard of Correctness). Ne voivat olla sopivia tai epäsoivia objektiensa suhteen, sillä ne ovat *representaationaalisia*. Toisin sanoen representaatio voi esittää kohteensa oikein tai väärin. Esimerkiksi uskomus p on sopiva, jos p on totta. Vastaavasti X:n pelko on sopiva, jos X voi vahingoittaa sinua. (Forrester 2024, 10–11.)

Haluilla on myös kohteensa. Ne representoivat kohteitaan tietynlaisiksi, eli hyviksi. Kysymys onkin, miksi uskomuksilla tai peloilla on jokin standardi, johon niiden oikeellisuutta voidaan verrata, mutta haluilla ei? Se, miten sopiva jokin asia on, ei ole ehdoton kyllä tai ei. Pikemminkin jonkin asian sopivuudelle on olemassa astevaihtelua. Jotkin halut voivat olla sopivampia kuin toiset. Vastaavasti mikään halu ei välttämättä ole täysin sopiva tai täysin sopimaton. (Forrester 2024, 11–12.)

Forrester ehdottaa, että tyydyttynyt halu tuottaa sitä enemmän hyvinvointia, mitä sopivampi se on. Vaikka halukone tyydyttää enemmän haluja, tyydytetyt halut eivät ole sopivia. Yksilön hyvinvointi on siis sopivuudella painotettu summa (engl. *fittingness-weighted sum*) tyydyttyneitä haluja. Halukoneessa ihminen esimerkiksi haluaa, että kaikki epäoikeudenmukaisuus maailmassa tapahtuu. Epäoikeudenmukaisuus ei voi lisätä ihmisen hyvinvointia, vaikka halu itsessään tyydyttyisi. Myös motivaatio asioiden muuttamiseen ja tavoitteiden asettaminen on hyvinvoinnin kannalta tärkeää, eikä niistä luopuminen ole sopivaa. (Forrester 2024.)

Olennaista ratkaisussa on se, että halujen puhtaasta subjektiivisesta tarkastelusta luovutaan osittain. Asioilla on haluttavuutensa, joka voidaan määritellä yksilön todellisista haluista riippumatta. Tästä huolimatta sopivat halut vaihtelevat eri yksilöiden välillä. Minun ei esimerkiksi ole sopivaa haluta koiranruokaa, vaikka koirani näkökulmasta sen haluaminen sitä on. Teoria on ”objektiivinen” siinä mielessä, että ulkopuolinen standardi, halujen sopivuus, osittain määrittää sitä, miten paljon hyvinvointia tyydyttynyt halu tuottaa. Toisaalta taas se, miten sopivia halut ovat, riippuvat subjektista. Haluteorian teorian tulisi hyödyntää tätä sopivuuden käsitettä, jotta halukoneen seuraukset voitaisiin

välttää. Toisin sanoen parempi teoria hyväksyisi tämän tyyppisen objektiivisuuden osaksi teoriaa. Vastalauseena voidaan Forresterin mukaan esittää, että intuitiomme halukonetta koskien ovat vinoutuneita, kuten myös kokemuskonetta on kritisoitu. Tästä huolimatta puhtaan subjektiivisuuden hylkääminen on ratkaisuna parempi. (Forrester 2024, 11–13.)

7.3. Kokemuskone: subjektiivisen ja objektiivisen välissä?

Tutkielmassani tarkasteltu kirjallisuus ei tietenkään ole täydellinen esitys kaikesta mitä kokemuskoneesta on kirjoitettu. Sen pohjalta voi kuitenkin tehdä joitain varovaisia johtopäätöksiä tai hypoteeseja: Kokemuskone ilmentää subjektiivisen ja objektiivisen, yksilön ja muun maailman, välistä rajapintaa. Hyvinvoinnin tai onnellisuuden kannalta omat kokemuksemme ja tuntemuksemme ovat keskeisessä asemassa. Subjektivistisen ajattelutavan äärimmilleen vieminen ei kuitenkaan vaikuta mielekkäältä tavalta toimia. Tämä ei kuitenkaan tarkoita, että subjektivismista tulisi kokonaan luopua.

Feldmanin ratkaisussa nautinnot tulee käsittää asenteellisiksi. Forresterin halukoneen ratkaisu sen sijaan vaatii, että haluja ei eroteta muista intentionaalisista mentaalista tiloista, kuten uskomuksista. Molempien ratkaisuihin siis käsitteet määritellään tietyllä tavalla, mikä on olennaista, jotta teoriat vastaisivat paremmin ajatuskokeiden ongelmiin. Feldmanin *totuuteen mukautettu hedonismi* vastaa suoraan kokemuskoneen ongelmaan. Forrester sen sijaan käyttää kokemuskonetta mallina omalle ajatuskokeelleen, halukoneelle. Ajatuskokeet eivät kummankaan kohdalla johda siihen, että kritisoitu hyvinvointiteoria hylättäisiin. Sen sijaan teorioita pyritään kehittämään ongelmien ratkaisemiseksi.

Feldmanin asenteellisten nautinnollisten jaksojen itseisarvo muodostuu siten, että nautinnon voimakkuus ja kesto kerrotaan totuuskertoimella. Forresterin teoriassa tyydyttyneiden halujen tuottama hyvinvointi on niiden sopivuudella painotettu summa. Se, miten sopivuuden painoarvo määritellään, jää avoimeksi. Feldman antaa valheellisten nautintojen painoarvoksi 0,1. Tämä on tietenkin lähes mielivaltaisesti valittu luku. Käytännön elämässä kukaan tuskin laskee nautintojensa itseisarvoja tai määritä niille painokertoimia. Sama pätee tyydyttyneiden halujen sopivuuden arviointiin. Tästä huolimatta on silti aiheellista kysyä: Millainen arvo valheellisilla nautinnoilla on suhteessa todellisiin nautintoihin? Entä miten sopiva jokin halu kullekin on?

Totuuteen mukautettu hedonismi ja sopivat halut ovat teorioita, jotka jossain määrin hylkäävät puhtaan subjektivistisen näkemyksen hyvinvoinnista. Ainakin on vaikea ajatella, että nautinnon totuus voitaisiin määritellä pelkästään subjektista käsin. Halujen sopivuuttakaan tuskin voi jättää täysin yksilön itsensä vastuulle. Yksi mielenkiintoinen seikka Forresterin artikkelissa on se, että siinä ennustetaan halukoneelle vastaavan kaltaista kritiikkiä, jota kokemuskone on saanut osakseen.

8. Yhteenveto: kokemuskoneen metodologisesta kiinnostavuudesta

Tässä kandidaatintutkielmassa olen tutkinut sitä, millä tavoin Robert Nozickin (1938–2002) kokemuskone (1974) edesauttaa filosofista teorianmuodostusta. Tarkasteluni on perustunut Kier Præem ja Asbjørn Steglich-Petersenin artikkeliin ”Philosophical thought experiments as heuristics for theory discovery” (2015). Ajatuskokeet voivat olla teorianmuodostuksen kannalta hyödyllisiä siten, että ne voivat vastaesimerkin kautta osoittaa teorian heikon kohdan, jonka ratkaisu vaatii teorian korjaamista. Ajatuskokeet voivat lisäksi tarjota valmiin mallin, jota voi hyödyntää etenkin muokkaamalla jo olemassa olevaa ajatuskoetta. Metodologisesti kiinnostavia tapauksia ovat ajatuskokeet, jotka jonkinasteisen säännöllisyyden puitteissa auttavat uusien hypoteesien ja teorioiden muodostamista.

Kokemuskone on vastaesimerkki hedonistiselle hyvinvointiteorialle. Hedonistisen teorian kannattajat ovat pyrkineet vastaamaan sen osoittamaan kritiikkiin. Sen seurauksena teoriaa on myös muotoiltu uudelleen. Esimerkkinä tästä olen käsitellyt Fred Feldmanin *totuuteen mukautettua hedonismia*. Tässä hedonismin muotoilussa valheelliset nautinnot erotetaan todellisista nautinnoista. Mikäli nautinnon kohde on epätodellinen, vähentää se myös *asenteellisen* nautinnon itseisarvoa. (Feldman 2004.)

Kokemuskonetta voi hyödyntää mallina myös muunlaisille filosofisille tarkasteluille. Tästä mallin hyödyntämisestä esimerkkinä olen käsitellyt Paul Forresterin *halukonetta*. Halukone on kokemuskoneen tarjoaman mallin pohjalta muodostettu ajatuskoe. Halukone on vastaesimerkki haluteorialla, vaikka kokemuskone ei sitä ole. Halukoneen ja sen mahdollisten ratkaisujen pohdinnan seurauksena Forrester ehdottaa, että haluteorian tulisi hyödyntää *sopivien halujen* -käsitettä. Halun kohteilla on haluttavuutensa, joka voidaan määritellä subjektista riippumatta. Kuten muillakin intentionaalisilla mentaalisisä tiloilla, myös haluilla on standardi, johon niiden oikeellisuutta voidaan verrata. Myös subjekti vaikuttaa siihen, miten sopiva halu on. (Forrester 2024.)

Muita tutkielman esimerkkejä, joissa kokemuskone on malli:

- (i) *Status quo* -vinouman ja *kuvitteellisen vastarinnan* testaaminen: Millaisia vaikutuksia erilaisilla kokemuskoneen muotoiluilla on ihmisten valintoihin. (De Brigard 2010; Weijers 2014; Hindriks & Douven 2017.) Tulisiko vinoumat huomioida filosofian opetuksessa? (Weijers 2014.)
- (ii) *Autenttisuuden intuition* testaaminen eri kokemuskoneen versioilla: ihmiset arvostavat elämää, joka on kosketuksessa todellisuuteen. Lisäksi ihmiset välittävät siitä, mitä he tekevät ja keitä he ovat. Arvostuksessa on astevaihteluita. (Hindriks & Douven 2017.)

(iii) Ovatko suoratoistopalvelu Netflixin yhä enemmän yksilön mieltymyksiä mukailevat ohjelmasuosituksiset eräänlaisia reaali maailman kokemuskoneita? (Brincker 2021.)

(iv) Skitsofreniaa sairastavan autonomia: Millaisin perustein muiden tulisi poistaa hänet vastentahtoisesti tällaisesta rikkinäisestä kokemuskoneesta? (Birks 2011.)

Utilitarismin tai *hedonismin* kritiikiksi tulkitulla kokemuskoneella voi olla muitakin käyttömahdollisuuksia. Sitä voi ja on myös käytetty erinäisten teorioiden ja hypoteesien muodostamisen apuna. Filosofian lisäksi kokemuskoneen tarjoama malli voi olla hyödyllinen myös empiiristen kysymysten tarkastelussa. Se, miten mielekkäitä tai kannatettavia teorioita, hypoteeseja ja näkemyksiä kokemuskoneen avulla voidaan muodostaa, on erillinen kysymys. Tästä huolimatta kokemuskone on teorianmuodostuksen kannalta metodologisesti kiinnostava ajatuskoe.

Lähteet

- BBC (2024). "Experience machines': The 1970s thought experiment that speaks to our times" <https://www.bbc.com/future/article/20240321-experience-machines-thought-experiment-that-inspired-matrixs-greatest-question> viitattu 13.12.2024.
- Berkovski, Sandy (2012). "The Possibility of Modified Hedonism", *Theoria* 78 (3), 186-212.
- Brendel, Elke (2018). "The Argument View: Are thought experiments mere picturesque arguments?", teoksessa Michael T. Stuart, Yiftach Fehige & James Robert Brown (toim.) *The Routledge Companion to Thought Experiments*. Lontoo: Routledge.
- Birks, David Rhys (2011). "Wellbeing, schizophrenia and experience machines", *Bioethics* 27 (2), 81-88.
- Bramble, Ben (2016). "A New Defense of Hedonism about Well-Being" *Ergo: An Open Access Journal of Philosophy* 3.
- Brincker, Maria (2021). "Disoriented and alone in the 'experience machine' - On Netflix, shared world deceptions and the consequences of deepening algorithmic personalization", *SATS* 22 (1), 75-96.
- Brown, James Robert & Fehige, Yiftach (2023). "Thought Experiments", teoksessa Edward N. Zalta & Uri Nodelman (toim): *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2023/entries/thought-experiment> viitattu 17.4.2024.
- Chen, M. Keith, Lakshminaryanan, Venkat, & Santos, Laurie (2006). "The evolution of our preferences: Evidence from capuchin monkey trading behavior", *Journal of Political Economy* 114, 517-537.
- De Brigard, Felipe (2010). "If You like It, Does It Matter If It's Real?" *Philosophical Psychology* 23 (1): 43-57.
- Feldman, Fred (2004). *Pleasure and the Good Life: Concerning the Nature Varieties and Plausibility of Hedonism*. Oxford: Clarendon Press.
- Feldman, Fred (2011). "What we learn from the experience machine", teoksessa Bader, Ralf M.; Meadowcroft, John (toim.): *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*, 59-86, 59.

- Forrester, Paul (2024). "The desire machine", *Analysis* 84 (2), 249-257.
- Hewitt, Sharon (2010). "What do our intuitions about the experience machine really tell us about hedonism?", *Philosophical Studies* 151 (3), 331–349.
- Hindriks, Frank, & Igor Douven (2017). "Nozick's Experience Machine: An Empirical Study", *Philosophical Psychology* 31 (2), 278–9.
- Horgan, Terence & Timmons, Mark (1992). "Troubles on moral twin earth: Moral queerness revived", *Synthese* 92, 221–260.
- Kawall, Jason (1999). "The Experience Machine and Mental State Theories of Well-being", *The Journal of Value Inquiry* 33, 381–387.
- Kolber, Adam J. (1994). "Mental Statism and the Experience Machine", *Bard Journal of Social Sciences* 3, 10-17.
- Lagerspetz, Eerik (2024). *Itsemääräminen ja valta*. Tallinna: Gaudeamus Oy.
- Lin, Eden (2016). "How to Use the Experience Machine", *Utilitas* 28 (3), 314-332.
- McGinn, Colin (1977). "Charity, interpretation, and belief", *Journal of Philosophy* 74, 521–535.
- Nozick, Robert (2008). *Anarchy, State and Utopia*. Blackwell publishing.
- Olsaretti, Serena (2007). "The Limits of Hedonism: Feldman on the Value of Attitudinal Pleasure", *Philosophical Studies* 136 (3), 409-415.
- Praëm, Sara Kier & Steglich-Petersen, Asbjørn (2015). "Philosophical thought experiments as heuristics for theory discovery", *Synthese* 192, 2827–2842.
- Sinnott-Armstrong, Walter (2023). "Consequentialism", teoksessa Edward N. Zalta & Uri Nodelman (toim): *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/win2023/entries/consequentialism/> viitattu 20.10.2024.
- Weijers, Dan (2014). "Nozick's experience machine is dead, long live the experience machine!", *Philosophical Psychology* 27(4), 513–535.
- Williamson, Timothy (2022). *Filosofinen menetelmä*. Suom. Tapani Kilpeläinen. Tallinna: Eurooppalaisen filosofian seura ry / *niin & näin*.