

Tuomo Kamula

MORAALINEN VASTUU SOTAROBOTIIKAN KÄYTÖSSÄ

Yhteiskuntatieteet
Kandidaatin tutkinto
12/2020

TIIVISTELMÄ

Tuomo Kamula: Moraalinen vastuu sotarobotiikan käytössä
Kandidaatin tutkinto
Tampereen yliopisto
Filosofian tutkinto-ohjelma
12/2020

Sotarobotiikkaa kehitetään yhä nopeammaksi ja sen ilmeinen kyvykkyys toimia itsenäisesti kasvaa. Tulevaisuudessa robotit saattavat olla kyvyiltään jopa ihmisen kaltaisia, jolloin kysymys siitä, kenellä on moraalinen vastuu sotarintamalla tehdyistä päätöksistä, nousee tärkeäksi moraaliseksi kysymykseksi. Tästä syystä on tarkasteltava, mistä syystä ihmisiä pidetään moraalisesti vastuullisena teoistaan, ja sitä pystyvätkö robotit täyttämään kyseisiä kriteerejä. Tarkastelen autonomisuuden käsitettä, milloin robotin voitaisiin nähdä olevan vastuussa omista teoistaan, ja tutkin myös, miten robottien käyttö sodankäynnissä vaikuttaa vastuun selvittämiseen.

Avainsanat: Sotarobotiikka, moraalinen vastuu, tahdonvapaus, autonomia, oikeutettu sota.

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

Sisällys

Johdanto	1
Robotiikan kehitys	2
Mitä on moraalinen vastuu?	3
Determinismi ja tahdonvapaus	4
Determinismi ja robotiikan tahdonvapaus	8
Autonomisen robotin vastuu	10
Käyttäjän vastuu.....	13
Pohdinta	19
Lähteet.....	21

Johdanto

Sotarobotiikkaa kehitetään ja tullaan kehittämään yhä nopeammaksi ja itsenäisemmäksi sodankäynnin välineeksi, ja tästä syystä on moraalisesti tärkeää selvittää, kuka on milloinkin vastuussa, kun sotarobotti tekee ihmishenkiä koskevia päätöksiä sotarintamalla. Perusteluna robotiikan kehittämiseksi usein annetaan, että sillä voi toteuttaa eettisemmin sodankäyntiä, ja säästää täten ihmishenkiä. Sotarobotiikan kehittäminen on kuitenkin myös kilpajuoksua, koska mitä nopeampia robotteja sodan osapuolet pystyvät kehittämään, sitä paremmassa asemassa he ovat sodankäynnin voittamisessa. Tulevaisuudessa robotiikan toiminta voi kiihtyä niin nopeaksi, että ihminen ei kykene enää olemaan mukana sen päätöksenteossa tekemättä siitä haavoittuvampaa.

Sotarobotiikka eli autonomiset asejärjestelmät tuovat oman haasteensa sodankäynnin eettisyyden suhteen, koska ne oletettavasti tulevat pystymään toimimaan autonomisesti eli tekemään päätöksiä itsenäisesti ennenäkemättömällä tavalla sotarintamalla. Nykyään armeijoilla on jo käytössä puoliautonomisia aseita, jotka ainakin ensivaikutelmaltaan toimivat osaksi itsenäisesti, kuten itseohjautuvat ohjukset. Robotiikan vastuukysymys on tärkeä, kun mietitään, onko ihminen vai robotti vastuusta korkean autonomisuuden tason omaavan robotin teoista.

Tässä tutkielmassa tarkastelen aluksi tahdonvapautta metafysiikan tasolla, joka auttaa selvittämään, mikä ominaisuus toimijalla on oltava, jotta hänen voidaan katsoa olevan moraalisesti vastuussa teoistaan. Tämän jälkeen tarkastelen kyseistä kysymystä myös robotiikan näkökulmasta, ja selvitän myös, milloin robotin voitaisiin nähdä olevan autonominen eli itsenäinen toimija. Tämän jälkeen tutkin oikeutettu sota -teorian avulla, miten robotiikan käyttö sodassa vaikuttaa käyttäjän moraaliseen kyvykkyyteen. Lopuksi selvitän, miten robotiikka vaikuttaa ihmishenkiä koskevien päätösten moraalisen vastuun selvittämiseen.

Robotiikan kehitys

Robotiikan kehityksen nähdään etenevän yhä enemmän suuntaan, jolloin ihmiset pystyvät suorittamaan sotatoimiaan suurempien välimatkojen päästä robotiikan avulla. Etäaseet ovat johtaneet sekä tehokkaampaan tappamiseen että vastapuolen vastarinnan välttämiseen. Sotilaat ovat monesti olleet sotien lähitaisteluissa haluttomia ampumaan, tai ovat tarkoituksella ampuneet yli ja ohi vihollisistaan. Mitä kauempaa sotatoimia on suoritettu, esimerkiksi tykistön pommituksen muodossa, sitä tehokkaampia ne ovat olleet vastapuolen tappamisessa. Tästä voitaisiin päätellä, että robotiikka tulee tekemään pitkän välimatkan aseena tappamisesta vähemmän psyykkisesti vaikuttavaa sen käyttäjälle. (Lin et al. 2011, 111-112.)

Muun muassa Lähi-idän konfliktialueilla on käytössä arviolta 6000–12000 maarobottia. Luku on kasvanut suuresti vuodesta 2004, jolloin se oli vain 150. Tämä osoittaa, että robotiikan kehittyessä ja sen käyttötapojen lisääntyessä robotteja otetaan myös enemmän käyttöön sotatoimiin. Tällä hetkellä maarobotit suorittavat paljon tylsiä, likaisia ja vaarallisia tehtäviä, kuten maamiinojen purkamista teiden reunoilta, tai vaarallisten alueiden tarkistamista vastapuolen taistelijoiden varalta. Ylivoimaisimmat sodankäynnin etäaseet tällä hetkellä ovat robottilentokoneet ja liidokit, joita ”pilotit” pystyvät lennättämään etänä pelikonsoleillaan, ja heidän tehtävänä on päättää, milloin käyttää tappavaa voimaa ja milloin ei. Roboteilla tehdyt iskut esimerkiksi Waziristanissa antavat osviittaa siihen, kuinka tuhoisia etäältä tehdyt hyökkäykset todella ovat. Lennokki-iskut ovat pakistanilaisen lehdistön mukaan tappaneen neljätoista Al-Qaeda-järjestön johtajaa, ja siinä sivussa yli 600 siviiliä (Lin et al. 2011, 114-115.)

Keskeinen agenda robotiikan kehityksessä on, että teknologiaa saadaan kehitettyä entistä nopeammaksi. Tämä koskee myös päätöksentekoa ampumisen suhteen, ja tavoitteena lopulta on mahdollisesti myös saada ihminen pois ampumisesta koskevasta päätöksenteon kehästä. Tämä tavoite on korkealla esimerkiksi Yhdysvaltojen armeijan tavoitelistalla. Ihminen työntyy ulos kehästä, koska robottien nähdään tulevan tekemään päätöksiä tulevaisuudessa mikro- ja nanosekunneissa. Tällöin ihmisen tehtäväksi jää enemmän valvoa päätöksenteon prosessia kuin olla mukana päätöksenteon kehässä. (Lin et al. 2011, 115-116.) Robotiikan nopeutuessa ja robottien autonomisten kykyjen kasvaessa ihmisen rooli esimerkiksi ampumisen päätöksenteossa ja kohteiden valinnassa supistuu, jolloin tärkeäksi muodostuu kysymys, kenellä on vastuuvastuu näiden robottien tekemistä päätöksistä sotarintamalla.

Mitä on moraalinen vastuu?

Galen Strawson määrittelee aidon moraalisen vastuun artikkelissaan *The Bounds of Freedom* seuraavasti: moraalinen vastuu on olemassa, jos ja vain jos toimijaa voidaan rangaista tai palkita reilusti tai oikeudenmukaisesti ilman, että rankaisulla tai palkitsemisella on mitään pragmaattista tai distributiivista oikeutusta (Strawson 2002, 452). Tällä siis tarkoitetaan sitä, että moraalisen vastuun määritelmän tulisi antaa objektiivinen oikeutus sille, että henkilö palkitaan tai tuomitaan moraalisesti toiminnastaan, ja että tämä palkitseminen ja tuomitseminen nähdään todella aidosti ansaittuna, eikä niitä perustella esimerkiksi pyrkimyksenä vaikuttaa henkilön käyttäytymiseen. Bruce Wallerin (2014) mukaan aidon moraalisen vastuun käsite kuitenkin määritellään usein tarkoituksella uudestaan, jolloin koko kysymyksenasettelu muuttuu: toimija on moraalisesti vastuussa, jos ja vain jos hänen rankaisemisellaan tai palkitsemisella on hyödyllinen vaikutus hänen luonteeseensa. Tämä ei Wallerin mukaan kuitenkaan vastaa alkuperäiseen kysymykseen, milloin rankaiseminen on aidosti reilua ja oikeudenmukaista, ja uusi määritelmä implikoi myös, että rangaistuksesta muuttumattomat, kuten paatuneet rikolliset eivät olisi moraalisesti vastuussa. Tällöin emme enää tarkastele, onko rankaiseminen aidosti oikeudenmukaista, vaan sitä, milloin se on pragmaattista. (Waller 2014, 11-13.)

Kun jollakin sanotaan olevan moraalinen vastuu, tarkoitetaan, että hänellä on tietynlainen yhteys tekoihinsa ja niiden seuraamuksiin. Henkilöön myös määritetään tietynlaisia ominaisuuksia ja kykyjä, joista hänen toimintansa nähdään oikealla tavalla johtuvan. Ensisilmäyksellä tarkasteltuna toimija on moraalisesti vastuussa, jos häntä kohtaan koetaan oikeutetuksi kohdistaa moraalista syylistämistä tai kehumista. Moraalinen syylistäminen ja kehuminen ovat epäsymmetrisiä ainakin sen huomion suhteen, miten niitä annetaan; syylistämiseen keskitytään paljon enemmän kuin kehunsaan. Kun jotakuta pidetään vastuussa jostakin teosta tai tapahtumasta, niin tämä ilmaisu ilmenee yleensä syyttämisen tapauksissa. Useimmat teokset moraalisen vastuusta keskittyvät vastuun negatiiviseen puoleen. (Matthew 2019.)

Vastuu-sanaa käytetään tarkoittamaan eri asioita eri konteksteissa, joka aiheuttaa ongelmia moraalisen vastuun määrittelyssä. Kaikkia näitä käyttötarkoituksia yhdistää tavalla tai toisella se, että henkilö, joka on vastuussa, on velvollinen antamaan selvityksen jostakin asiaintilasta. Toisaalta taas henkilö voi olla myös joidenkin velvoitteiden alaisena, jotka asettavat tietyt standardit hänen toiminnalleen, jolloin hän on myös altis muiden reaktioille häntä kohtaan. Jos otetaan esimerkkinä vaikkapa laivan kapteeni, jolla on tärkeä velvollisuus hänen miehistöään ja matkustajia kohtaan: hänen pitää varmistaa, että he selviävät turvallisesti matkasta. Tällöin voitaisiin sanoa, että kapteenilla

on vastuu laivan kyydissä olevien turvallisuudesta, tai että henkilöiden turvallisuus on kapteenin vastuulla. (Talbert 2016, 12.)

Determinismi ja tahdonvapaus

Rikoshistoriasta tunnetusta tapauksesta, jossa rankan lapsuuden kärsinyt Robert Harris tappoi kaksi nuorta poikaa väkivaltaisesti, on helppo tehdä arvostelma, että hän on paha ja julma ihminen. Kuitenkin kun arvioidaan, että onko Harris moraalisesti vastuussa teoistaan, hänen rankka lapsuutensa tuo mukanaan syitä sille, miksi Harriksesta tuli luonteeltaan julma, eikä hänen lapsuutensa kärsimyksen, kuten väkivaltaisen isän, voida nähdä olleen hänen hallinnassaan. Tällöin vaikka katsottaisiin, että Harrista voidaan syyttää teoistaan, koska hänen julmaan luonteeseensa kuuluu, että hän tekee julmia tekoja, niin avoimeksi jää kysymys, että ansaitseeko hän todella moraaliset syytökset ottaen huomioon hänen lapsuudestaan juontuvat hallitsemattomissa olevat tekijät. (Waller 2014, 22.)

Jotta toimijan voitaisiin aidosti sanoa olevan vastuussa teoistaan, hänelle pitäisi olla mahdollisuus toimia toisin, kuin hän päätti toimia. Tämä tarkoittaa, että toimijalla pitää olla vapaa tahto valita tekonsa. Vapaan tahdon mahdollisuuden asettaa kyseenalaiseksi kausaalinen determinismi, jonka mukaan kaikki tapahtumat ja teot aiheutuvat aikaisemmista tapahtumista ja teoista luonnonlakien mukaisesti, ja tämä aiheuttaisi sen, että tulevaisuus ja nykyhetki ovat ennalta määrättyjä, ja voivat tapahtua vain yhdellä ainoalla tavalla. Tahdonvapauden nähdään täten usein olevan yhteensopimaton kausaalisen determinismin kanssa. (Matthew, 2019.) Ainoastaan silloin, jos meillä olisi ihmisinä mahdollisuus rikkoa luonnonlakeja, tai muuttaa menneisyyttä, voitaisiin nähdä, että tahtomme on vapaa muuten deterministisessä maailmassa (Talbert 2016, 17).

Näkemyistä tahdonvapauden yhteensopimattomuudesta determinismin kanssa kutsutaan inkompatibilismiksi. Inkompatibilistit voivat joko katsoa, että tahdonvapaus on olemassa ja determinismi on virheellinen, tai sitten he voivat katsoa, että tahdonvapautta ei ole. Ensimmäistä ryhmää kutsutaan libertarianisteiksi ja jälkimmäistä vapaan tahdon skeptikoiksi. Kolmas vastaus tahdonvapauden ongelmaan on kompatibilismi, eli determinismin yhteensopivuus tahdonvapauden kanssa. (Talbert 2016, 18.)

Kompatibilistit uskovat, että determinismin hyväksyminen ei tarvitse johtaa siihen näkemykseen, että vaikka toimija tekisi mitä tahansa, hänen kohtalonsa johtaisi kuitenkin aina tiettyyn ja samaan

lopputulokseen. Päätöksien teosta voidaan silti katsoa, että valitsemamme toiminta johtaa tekemiemme päätösten kautta eri lopputulemiin. (Talbert 2016, 18-19.) Esimerkiksi henkilö X olisi voinut valita tienhaarassa reitin A sijasta reitin B. Vapaan tahdon skeptikot vastaavat tähän vastaväitteeseen kuitenkin sanomalla, että tahdonvapaudelle aidosti relevanttia on, että henkilön X päätöksentekoa ei saisi ohjailta aiemmat tapahtumat ja luonnonlait. Henkilöllä X voi kyllä ilmetä, että hän puntaroi mahdollisia vaihtoehtoja mielessään, ja hän kokee tekevänsä päätöksiä, mutta vapaan tahdon skeptikot näkevät tämän päätöksen tapahtuvan kuitenkin determinoidusti. (Talbert 2016, 19.) Daniel Dennett on muun muassa antanut monta eri kompatibilistista muotoilua tahdonvapaudelle, muttei kuitenkaan kirjassaan *Freedom Evolves* saavuttanut yhtäkään itseään tyydyttävää. Lopulta hän päätyi toteamukseen, että koska ilmiselvästi olemme moraalisesti vastuussa, tämän takia mitä ikinä valtaa ihmisillä onkaan, niin sen on riittävä moraaliseen vastuuseen. (Dennett 2003, 297).

Kompatibilismista on olemassa myös utilitarianistinen versio, jossa tosiaan katsotaan, että jos toimijaa on hyödyllistä rangaista tai palkita teoistaan, niin silloin hänen voidaan nähdä olevan vapaa toimintansa suhteen. Utilitarianistit hyväksyvät, että mitään tapahtumaa ei ole ilman kausaalista aiheuttajaa, mutta he tekevät erottelun rajoitettujen (constrained) ja vapaiden tekojen välillä. Esimerkiksi kleptomaanikkoa tai mielisairasta ei olisi järkevää yrittää muuttaa uhkailemalla tai palkitsemisella. Samaa pätee esimerkiksi tekoihin, jotka on tehty pakotuksen alla tai stressaantuneena. Käänteisesti taas, jos rankaisulla tai kehuilla pystytään ohjailemaan toimijan käytöstä, hänen tahtonsa on vapaata. (McFee 2000, 70-71.)

Tässä vaiheessa on tehtävä selvä erottelu kahden erityyppisen tahdonvapauden määritelmän välillä, sillä vaikuttaisi siltä, että kompatibilistit ja kovan determinismin kannattajat puhuvat hiukan eri asioista. David Hume on muun muassa määritellyt tahdonvapauden kykynä tehdä mitä itse haluaa, tai toimia omien mielihalujen pohjalta, ja tämän hän sanoo olevan yhteensopiva determinismin kanssa. (Balauger 2014, 49.) Kovan determinismin kannattajat vastaisivat tähän sanomalla, että kyseiset mielihalut ovat kausaalisesti determinoituja henkilön syntymää edeltämistä tapahtumista, eikä hänen siksi voida nähdä olleen vapaa päättämään, mitä haluaa. Toisentyyppinen tahdonvapauden määritelmä olisi, että toimijan päätökset eivät saa olla samaan aikaan sekä ennalta determinoituja että myöskään täysin sattumanvaraisia ollakseen aidosti vapaita. (Balauger 2014, 50). Tämän määritelmän valossa kompatibilistinen määritelmä on irrelevantti muun muassa Immanuel Kantin ja William Jamesin mielestä, sillä heidän mukaansa kompatibilistit väistävät oikean ongelman siirtämällä keskustelun determinismin ja tahdonvapauden yhteensopimattomuudesta siihen, että toimijan tunne, että hän tekee päätöksiä, riittäisi tahdonvapauteen (Balauger 2014, 54).

Libertarianisteilla on kaksi mahdollista tapaa kieltää determinismin oletus, että kaikki tapahtumat ovat ennalta kausaalisesti määrättyjä. Ensimmäinen tapa on kieltää, että joillakin tapahtumilla ei ole kausaalista syytä, ja toinen on, että millään tapahtumalla ei ole kausaalista syytä (McFee 2000, 53.) Esimerkiksi Charles Arthur Campbell esittää libertarianistisen näkemyksen, että vaikka muut tapahtumat ovat kausaalisesti determinoituja, niin kausaalisessa linkityksessä on epäjatkuoita liittyen ihmisen moraaliseen toimintaan. Hänen mukaansa, jos teot tehdään taipumuksesta (inclination) niin ne ovat kausaalisesti determinoituja, mutta jos ne tehdään velvollisuudesta (duty), niin tällöin kausaalinen jatkuo ei säily. Hänen mukaansa introspektio tekee tästä ilmeisen selvää, sillä toimiessamme velvollisuudesta, taistelemme vahvimpia halujamme ja himojamme vastaan, jotka toimivat kausaalisena voimana, ja tunnemme tekevämme päätöksiä. Samalla yleensä velvollisuudesta toimiminen pitää sisällään moraalisen velvoitteen (obligation) ja täten Campbell näkisi, että moraaliset teot ovat epädeterminoituja. (McFee, 2000, 55.)

Deterministi vastaa Campbellin argumenttiin toteamalla jälleen, että toimija voi introspektion avulla kyllä tuntea olevansa vastuussa teoistaan, mutta esimerkiksi kleptomaniikko kokee samalla tavoin, että hän olisi vastuussa teoistaan, mutta silti koska tunnemme hänen mielenhäiriönsä, emme pidä häntä aidosti vastuullisena. (McFee 2000, 56.) Samalla tavoin myöskään kukaan meistä ei ole valinnut biologista perustaamme, emmekä täten pysty päättämään, mitä tekoja koemme haluavamme tehdä. Toiseksi Campbellin tekemä jako velvollisuudesta ja halusta tehtyjen tekojen välillä ei ole ihan niin selkeä kuin Campbell tahtoo uskoa. (McFee 2000, 56.) Esimerkiksi hyväntekijä voi tehdä lahjoituksia, ja luultavasti Campbell näkisi tämän johtuvan hänen velvollisuudentunteestaan, mutta se voi johtuakin henkilön taipumuksesta tehdä hyviä tekoja muiden ihmisten puolesta. (McFee 2000, 56.) Campbellin teoriasta voidaan myös sanoa, että se ei pelastaisi tarpeeksi suurta osaa ihmisten toiminnasta kausaaliselta kehältä, sillä vain teot, jotka perustuvat velvollisuudelle, on hänen mukaansa vapaita. Silti jos haluamme nähdä olevamme vapaita toimijoita, kaikki tekomme pitäisi olla vapaita, ja näihin lukeutuu myös huoneen läpi käveleminen tai ruoan syöminen, joiden ei oletettavasti nähdä perustuvan velvollisuuteen. (McFee 2000, 57.)

Libertarianistisissa teorioissa on jaottelu tapahtumakausaalisuuden ja toimijakausaalisuuden välillä. Toimijakausaalisuudessa tapahtuman aiheuttajana pidetään substansseja, joiden määritellään olevan yksittäisiä olioita tai entiteettejä. Täten olisi toimijakausaalisuuden mukaista sanoa vaikka, että koira söi pöydältä ruokaa, tai tyttö rikkoi ikkunan. Substansseja näissä tapauksissa ovat koira ja tyttö. Oletuksena voidaan pitää, että objektit, eläimet, henkilöt, ja asiat kuten vesi ja happo ovat substansseja. Tähän ryhmään ei kuulu esimerkiksi objektien ominaisuudet, kuten keltaisuus tai

pyöreys. Tapahtumakausaalisuudessa tapahtuman aiheuttajana pidetään toista tapahtumaa, esimerkiksi sähköjohdon kipinäointia, joka sytyttää metsäpalon. (Griffith 2013, 80.)

Tahdonvapauden kannalta tämä erottelu tarkoittaa sitä, että esimerkiksi Thomas Reidin mukaan ihmisellä on ainoastaan substanssina kausaalista voimaa, sillä ihminen pystyy tekemään asioita, toisin kuin tapahtumat, jotka vain tapahtuvat. Aidot kausaaliset voimat tarvitsevat hänen mukaansa mielen ja älykkyyden, mitä ainoastaan ihmisillä on. Muut tapahtumat, jotka eivät ole ihmisen aiheuttamia, kuten ukkoset, aiheuttaa Jumala. Muunlaisen toimijakausaalisuuden teorian on muotoillut Roderick Chisholm, ja hänen mukaansa maailmassa on olemassa kahden tyyppisiä tapahtumien aiheuttajia. Toinen niistä on vapaa ja intentiosta toimiva ihminen, ja loput ovat tapahtumia. Chisholmin mukaan tapauksessa, jossa esimerkiksi mies liikuttaa sauvaa, ja tämä kyseinen sauva liikuttaa siinä olevaa jalokiveä, mies laittaa toimijakausaalisuudellaan neurobiologiset tapahtumat käyntiin, joka johtavat tapahtumakausaalisuusketjua pitkin liikuttamaan jalokiveä. (Griffith 2013, 82.)

Toimijakausaalisuuden kannattajien mukaan mikään muu ei determiinoidu toimintaa, kuin toimija itse, ja täten he kuvailevat toimijaa nimeltä ”determiinoidun determiinoidun”. Yhteensopimaton teorian kanssa siis on esimerkiksi havaittu aivotapahtumien yhtäläisyys toimintaan johtamisen kanssa, sillä nekään eivät voi teorian mukaan olla mukana vaikuttamassa agentin toimintaan. Teorian kriitikot pitävät sitä liian mystisenä, sillä esimerkiksi tiede ei tukisi sitä. Toiset taas katsovat, että teoria ei anna mitään selitystä sille, miksi toimija valitsee lopulta jonkin vaihtoehdon toiminnalleen jonkun toisen toiminnan sijaan. (Griffith 2013, 85.)

Kova determinismi on siis näkemys, jonka mukaan kaikki tapahtumat maailmassa ovat ennalta määrättyjä, ja jos tietäisimme kaiken aikaisemmista tapahtumista ja luonnon lainalaisuuksista, osaisimme ennustaa tulevaisuuteen. Kovassa determinismissä nähdään, että yksilö ei ole voinut vaikuttaa ennen syntymäänsä ilmeneviin tapahtumiin tai luonnonlakeihin. Saul Smilansky kirjoittaa artikkelissaan *Hard determinism and punishment: A practical “reductio”*, kovan determinismin kannan johtavan siihen, että yksilön tahto ei ole vapaa, eikä häntä myöskään voida pitää moraalisesti vastuussa toiminnastaan. Hänen mukaansa monesti tästä syystä kovan determinismin kannattajat ovat puolustaneet ajatusta, että meidän ei tulisi rangaista pahantekijöitä retributiivisista- eli kostamissyistä. (Smilansky 2011, 354.)

Pragmaattisista syistä kovan determinismin kannattaja voi suosia näkemystä, että rikolliset lukitaan ulos yleisestä väestöstä, eikä toisin päin, sillä yhteiskunta, jota kontrolloisi rikolliset, ei olisi moraalisesti hyväksyttävä. Kuitenkin vaikka pahantekijät lukittaisiinkin pois yleisestä väestöstä, niin koska kovan determinismin mukaan kukaan ei pohjimmiltaan ansaitse aidosti moraalista syytöstä

teoistaan, rikollisten pitää saada kompensatiota tästä eristyksestä. Heidän olonsa tulisi tehdä mahdollisimman mukavaksi, ja mukavuuden taso tulisi suhteuttaa muiden yhteiskunnan sosiaalisten tarpeiden kulujen ja taloudellisten rajoitteiden mukaan. Muut sosiaaliset tarpeet eivät saisi kärsiä liikaa rikollisten mukavuuden edessä. (Smilansky 2011, 355-356.)

Kovaa determinismiiä seurailleen rankaisemisen hyötyperuste heikentyisi, sillä rikoksien tekemisestä tulisi voitto-voitto-tilanne, kun tekijöitä ei enää uhkaisi rangaistus, vaan mahdollisimman mukavat oltavat jäädessään kiinni. Monelle tämä saattaisi olla myös rikollisuuteen houkuttava tekijä, sillä kiinnijääneille pitäisi kompensoida heidän vapautensa riistäminen muilla tavoilla, ja nämä kompensatiot voisivat aiheuttaa positiivista houkutusta verrattuna yhteiskunnassa elämiseen. Huomattavasti useampi ihminen sekaantuisi rikollisuuteen, ja yhteiskunta ajautuisi suurien rikosaaltojen alle, samalla kun rikollisten ylläpidosta koituisi kohtuuttomia kustannuksia. (Smilansky 2011, 359-360.)

Koska moraalisen syytöksen tai ylistyksen aitoon ansaitsemiseen ei ole olemassa mahdollisuutta kovan determinismin mukaisessa maailmassa, vaihtoehtona tietenkin on, että se unohdettaisiin kokonaan, ja noudatettaisiin jotakin pragmaattista moraaliteoriaa. Aito ansaitseminen on kuitenkin tärkeä käsite, kun tarkastellaan, mikä on moraalista perustilanteesta poikkeavaa. Yleisesti voidaan katsoa, että moraalissa perustilanteessa kaikkien nähdään olevan syyttömiä, kunnes toisin todistetaan, ja jos kova determinismi pitää paikkansa, niin kukaan ei myöskään ole missään tapauksessa oikeudenmukaisesti syyllinen. (Smilansky 2011, 364.)

Moraali on keskeisesti käytännöllistä, joten moraaliteorialle on tärkeää olla yhteydessä sen käytännöllisyyden kanssa. Moraalisuus auttaa ihmisiä elämään toistensa kanssa, ja välttämään toistensa vahingoittamista tässä prosessissa. Mahdollista on kuitenkin nähdä, että moraalinen teoria voi pitää paikkansa, vaikka sitä ei pystykään ottamaan käytäntöön. Näin saattaa olla kovan determinismin tapauksessa, ja tämä aiheuttaa sen, että moraaliteoria ja käytäntö ajautuvat toisistaan erilleen. (Smilansky 2011, 366-367.) Seuraavassa kappaleessa käsittelen aihetta robotiikan tapauksessa.

Determinismi ja robotiikan tahdonvapaus

Jatkuva teknologian kehittyminen robotiikassa ja tietoteknologiassa johtaa yhä tehokkaampaan tehtävien laskennalliseen voimaan, ja tämä kehitys saattaa tuottaa tulevaisuudessa lähes ihmisen -, ihmisenkaltaisen- tai ihmistä korkeammat älykkyyden kyvyt. Visiot siitä, että roboteista voisi jossain

vaiheessa tulla rationaalisia, itsetietoisia tai tietoisia, ovat johtaneet pohdintoihin robottien oikeuksista ja laeista, jotka käsittelevät ihmisen ja robotin yhteiskunnallisia asemia ja suhteita. Ihmisyhteisössä oikeuksia tasapainottavat velvollisuudet ja vastuut. (Ashrafian 2015, 2.)

Jos robottien moraalista vastuuta tarkastellaan determinismin näkökulmasta, niin ihmisiä on muun muassa verrattu determinismin takia robotteihin, koska heidän toimintansa nähdään olevan luonnonlakien ja aikaisempien tapahtumien ennalta määräämä. Kovan determinismin näkökulmasta robotit pystyvät tekemään tekoja omasta vapaasta tahdostaan vain, jos niiden toimintaa ei ole ennalta ohjelmoitu niihin. Libertarianistisesta näkökulmasta roboteilla on vapaa tahto, ja ne myös täten voivat olla aidossa mielessä moraalisesti vastuussa toiminnastaan. Kompatibilismi taas näkisi asian niin, että roboteille on ohjelmoitu rationaalisuus ja aistimuksellisuus, ja ne näiden kautta pystyvät saavuttamaan vapaan tahdon. (Ashrafian 2015, 3.)

Keskeinen robotiikan vastuuta koskeva kysymys on, että voidaanko sillä nähdä olevan vapaata tahtoa vai ei. Esimerkiksi kovan determinismin edustajat siis näkevät samoin robottien kuin ihmisten suhteen, että robottien teot ovat fysiikan lakien ennalta määrättäviä, eikä täten roboteilla ole vapaata tahtoa, eikä myöskään aitoa vastuuta teoistaan. Libertarianismin mukaan taas ihminen on esimerkiksi vapaa toimija, ja samalla lailla vaikkapa robotin sielu tai järjellisyys voisi tehdä siitä vapaan. (Ashrafian 2015, 4: Griffith 2013, 82.)

Harry Frankfurt on luonut ajatuskokeen, jossa valinnan vapauden puutteen ja moraalisen vastuun yhteensopimattomuus asetetaan kyseenalaiseksi. Yksilön voidaan siis nähdä olevan vastuussa teoistaan, vaikka hänellä ei olisikaan oikeasti mahdollisuutta valita toisin. Ashrafian muotoilee kyseisen ajatuskokeen sotarobotiikkaan soveltuvaksi:

Sotarobotti A on todennäköisesti suorittamassa tehtävän, jolla se varmistaa voiton sodasta omalle maalleen, samalla kuitenkin kiinnittäen myös huomiota paikallisten asukkaiden hyvinvointiin kansainvälisiä sopimuksia noudattaen. Ainoastaan tapauksissa, jossa asukkaista on suora uhka robotin oman maan sotilaita tai sotarobotteja kohtaan, se jättää paikallisten hyvinvoinnin huomioonottamatta. Robotti A:n ohjelmoija Y kuitenkin haluaa, että robotti A toimii kaikissa mahdollisissa tapauksissa paikallisten asukkaiden hyvinvoinnin eduksi, ja asettaa sen virtapiireihin erikoisimplantin, joka ohittaa robotti A:n ohjelmoinnin, jos implantti rekisteröi, että robotti A meinaisi toimia vastoin paikallisten hyvinvointia. Robotti A kuitenkin toimii omasta aloitteestaan paikallisten hyväksi, jolloin implantin ei tarvitse vaikuttaa robotin toimintaan.

Tällöin vaikka robotilla ei nähtäisi olevan mahdollisuutta valita toisin, se kuitenkin nähtäisiin intuitiivisesti olevan vastuussa toiminnastaan. (Ashrafian 2015, 4.)

Frankfurtin esimerkit tukevat kompatibilistista näkemystä, jonka mukaan determinismi on yhteensopiva tahdonvapauden kanssa, ja niillä pyritään myös osoittamaan, että inkompabilitismi olisi epätosi. Täten, vaikka robotit ovatkin ohjelmoituja ja keinotekoisia, ne ovat silti aistivia (sentient) ja moraalisesti vastuussa teoistaan. (Ashrafian 2015, 5.) Myöskään Frankfurtin esimerkit eivät selviä täysin ilman kriittisiä haasteita, sillä esimerkiksi kyseenalaiseksi jää, onko robotilla todella vain yksi valinnan mahdollisuus, koska robotti pystyisi toimimaan sekä omasta tahdostaan, että implantin vaikuttamana. (Matthew 2019.) Toisekseen avoimeksi jää, miten implanti voi olla varma, että robotti meinaa toimia paikallisten asukkaiden hyvinvoinnin eduksi. Jos tämä varmuus selitetään sillä, että implanti havaitsee joitakin kausaalisia tekijöitä, jotka varmistavat robotin toimivan juuri paikallisten hyödyksi, niin tällöin kysymykseksi jää, onko esimerkin maailma deterministinen. Tässä tapauksessa inkompatibilistit eivät tietenkään salli, että robotti olisi moraalisesti vastuussa kyseisessä maailmassa. (Matthew 2019.)

Autonomisen robotin vastuu

Tutkiaksemme sitä, että ovatko robotit missään tapauksissa vastuussa teoistaan, täytyy meidän tehdä erottelu kausaalisen ja moraalisen vastuun välille. Usein on selvää, että henkilö on kausaalisesti vastuussa, jos hän on aiheuttaja tai keskeinen (salient) syy jollekin tapahtumalle. Kausaalisen vastuuseen vaadittavat kyvyt eivät ole kuitenkaan yhtenevät moraalisen vastuuseen vaadittavien kykyjen kanssa, joten vaikka yksilö voitaisiinkin nähdä olevan kausaalisesti vastuussa, tämä ei suoraan tarkoita, että hän olisi myös moraalisesti vastuussa. (Matthew 2019.) Loogikot ovat tehneet toiminnalle kaksi eriävää määrittelyä, ja ne ovat Chellasin *Seeing to it that*, ja tämä on teoria kausaalisesta vastuusta, ja toinen määrittely, joka taas on Hortyn *Deliberately seeing to it that*, koskettaa moraalista vastuuta, sillä tässä määrittelyssä *deliberately*-sanalla viitataan siihen, että henkilö on kyennyt tarkoituksenmukaisesti tekemään valinnan toimintansa suhteen. (Lin et al. 2011, 149-151.)

Esimerkiksi tapauksessa, jossa sotajoukon komentaja määrää robotille jonkin tehtävän, niin tällöin kyseistä toimintaa voitaisiin kuvata näin: komentaja A tarkoituksella (*deliberately*) katsoo, että robotti C pitää huolen siitä, että P. Molempien voidaan nähdä olevan kausaalisesti vastuussa tapahtumasta P, mutta ainoastaan komentajan nähdään olevan moraalisesti vastuussa, sillä hän olisi voinut valita

toimintansa toisin, mitä mahdollisuutta robotilla ei ole. Kyseisissä määritelmässä katsotaan, että toimijalla täytyy olla mahdollisuus toimia toisin, jotta hän voisi olla moraalisesti vastuussa, ja että robotilla ei ollut mahdollisuutta toimia toisin. (Lin et al. 2011, 151.)

Ajatukset robotin ja komentajan moraalisen vastuun tasoista on yhteneväinen kahden lainsäädännön maksiimin kanssa. Toinen niistä on nimeltään *Qui facit per alium facit per se*, ja tämä maksiimi tulkitaan siten, että hän, joka toimii toisen kautta, tekee teon itse. Kyseistä maksiimia käytetään yleensä selittämään, miten työnantaja palkatessaan työntekijän tekemään tietyn tehtävän, on työnantaja vastuussa tämän tehtävän teosta niin kuin hän olisi itse sen tehnyt. Toinen maksiimi lainsäädännössä on nimeltään *Respondeat superior*, joka on hyvin samanlainen kuin aikaisempi, ja se tulkitaan ”herran” vastuuksi vastata. Sääntöä kutsutaan myös herra-palvelija-säännöksi, jonka mukaan useimmissa tapauksissa työnantaja on vastuussa työllistämänsä työntekijän tekemistä teoista. (Lin et al. 2011, 151.)

Moraalifilosofiassa toimijan autonomialla tarkoitetaan yleensä mahdollisuutta toimia itse, kykyä ohjailta itse omaa toimintaansa ja lopulta myös moraalista vastuullisuutta omasta toiminnastaan. Laiteautonomian suhteen on kuitenkin olemassa erinäisiä käsityksiä siitä, miten se määritellään, ja tämä tietysti vaikuttaa myös kysymyksiin laitteen moraalista vastuusta. Muun muassa sellaisen laitteen toimintaa, joka ei vaadi ihmisen minkäänlaista väliin tulemistä suoriutuakseen tehtävistään, luonnehditaan usein autonomiseksi. Vaikka myös ihmisen mahdollisuus olla autonominen toimija deterministisessä maailmassa on filosofiassa edelleen väittelyn alaisena, yhteisymmärrystä herättää kuitenkin ajatus, että robotit kehittyvät koko ajan autonomisemmiksi, jolloin niiden korkeampi autonomisuuden taso tekee niistä vähemmän ymmärrettäviä ihmisille, ja tämä johtaa ihmisen vähenevään kontrolliin. (Noorman & Johnson 2014, 52-53.)

Esimerkiksi Robert Sparrow on yrittänyt määritellä autonomiaa robottikysymyksessä siten, että toimija on autonominen, jos hänen tekonsa ovat hänestä lähtöisin ja ne refleктоivat hänen tavoitteitaan. Täysin autonomiset toimijat ovat hänen mukaansa myös jollain tavalla valinneet nämä tavoitteet itselleen. Kyseisessä määritelmässä Sparrow jättää kuitenkin avoimeksi sen, mitä hän tarkoittaa toimijan tavoitteilla ja niiden valitsemisella. Sparrow katsoo, että ketään ihmistä ei voida tämän määritelmän mukaan pitää vastuussa robottien toiminnasta, sillä robottien toimiminen omien halujen, toiveiden ja tavoitteiden pohjalta tekee niistä arvaamattomia, ja tämä tekee moraalisesti vastuussa olemisen osoittamisesta ongelmallista. Robottien luotettavuus ja hallittavuus riippuu siitä, kuinka hyvin tiedämme miten ne toimivat. Autonomisten robottien todennäköisen toiminnan ennustamisen suhteen tulisi meidän esimerkiksi tietää, ovatko niiden toiminnat determinoituja vai

epädeterminoituja, millaisia rajoitteita niiden toimintaan voi ja pitäisi ohjelmoida, ja mitä rajoitteita robotiikan luotettavuuden ja ennustettavuuden testauksessa on. (Noorman & Johnson 2014, 54.)

Lopulta päätös siitä, käyttävätkö ihmiset robotteja, jotka eivät ole riittävän luotettavia tai ennustettavia, on ihmisten harteilla. Tähän vaikuttavat vastuutoimenpiteet, jotka ihmiset luovat robottien kehittämistä varten. Niiden avulla päätetään, kuinka robottien halutaan operoivan, milloin niihin luotetaan tarpeeksi, että ne suorittavat ohjatut tehtävänsä, miten ne otetaan käyttöön, kuinka tarkasti niiden toimintaa pitää valvoa ja niin edelleen. Jos huomio kiinnitetään robotteihin, joiden toiminnasta kukaan ihminen ei pysty olemaan vastuussa, niin tällöin vähemmälle huomiolle jää ihmisten omaa vastuu robotiikan kehittämisessä ja käyttöönotossa. (Noorman & Johnson 2014, 54-55.)

Robottien saatetaan nähdä myös tulevaisuudessa olevan vastuussa teoistaan kasvavan autonomisuutensa ansiosta. Tällöin robottien autonomisuutta verrataan ihmisten autonomisuuteen, ja kun roboteista tulee toimintansa autonomisuudessa yhä enemmän ihmisen kaltaisia, saatetaan niitä alkaa harkita jo esimerkiksi lain edessä vastuullisiksi. (Noorman & Johnson 2014, 55.) Tätä teoriaa varten on keskeistä ajatella, että autonomisuus on eräänlainen jatkumo, ja eri teknologiat asuttavat tämän jatkumon eri kohtia riippuen siitä, mitä ominaisuuksia niillä on. Peter Asaro (2008) luettelee artikkelissaan *How Just Could a Robot War Be?* autonomisuudelle merkityksellisiä ominaisuuksia, joita ovat muun muassa kyky aistia, päättää ja toimia ilman ihmisen väliin tulemista. (Asaro 2008, 2.) Asaro lisää myös, että robotit saattavat tulevaisuudessa pystyä saavuttamaan omatakeisen moraalisen kyvykkyyden, jolloin ne pystyisivät itse muodostamaan muun muassa omia moraalisia periaatteita, velvollisuuksia ja syytä toiminnalle. Asaron visioimat autonomisuudelle keskeiset kyvyt ovat kuitenkin metaforisesti ilmaistuna, sillä aistimista, ajattelua ja toimimista käytetään yleensä kuvailemaan ihmisen toimintaa, ja roboteissa nämä toiminnat saattavat aktualisoitua merkittävästi erilaisella tavalla. (Noorman & Johnson 2014, 55.)

Koneautonomia on monisyinen ja mielipiteitä jakava konsepti. Koneiden autonomian käsitettä käyttävät eri tutkijat eri tavoilla, jolloin päätelmissä saatetaan päätyä erinäisiin lopputulemiin siitä, voivatko koneet ikinä olla moraalisesti vastuussa, tai aidosti autonomisia toimijoita. Muun muassa Yhdysvaltojen puolustusvoimilla on ollut merkittävä rooli robottiautonomisuuden käsitteen määrittämisessä, sillä puolustusvoimien jaostot perustuvat vahvasti kontrolliin ja hallintaan, ja robottien odotetaan täten olevan samoin rajoitteiden sisällä toimivia ja ennustettavia teknisiä systeemejä. Yhdysvaltojen puolustusvoimien tutkimus- ja kehitysjaosto on asettanut yhdeksi viidestä päätavoitteestaan autonomiset kulkuneuvot. On erilaisia näkemyksiä siitä, työntyykö ihminen aseellisten robottijärjestelmien päätöksenteon kehästä ulos niiden kasvavan nopeuden takia. Yksi

näkemyks on, että laitteiden prosessointikyky tulee olemaan niin nopea, että ihminen hidastaisi niiden toimintaa ja tekisi siitä haavoittuvaisemman. Toinen näkemys on, että koneita kehitetään nykyhetken vahvasti ihmisen vaikutuksen alaisena olemisesta autonomisempaan suuntaan, jolloin ihmiset pystyisivät kuitenkin harjoittamaan enemmän oikea-aikaista ja informoitua valvontaa. (Noorman & Johnson 2014, 56.)

Heather Roff (2014) olettaa artikkelissaan *The Strategic Robot Problem: Lethal Autonomous Weapons in War* robottiautonomisuuden sellaisille laitteille, jotka kykenevät oppimaan ja ottamaan kohteeksi kohteita, jotka eivät lähetä signaaleja, tai kulje tietyillä liikeradoilla tai nopeuksilla. Esimerkiksi Yhdysvaltojen MK – 15 Phalanx -järjestelmä toimii tällaisten rajoitteiden sisällä. Se pystyy automaattisesti kohdistamaan tietynlaisia ohjuksia ja lentoaluksia, ja tuhoamaan niitä, mutta sen toiminta ei pysty ulottumaan sen ennalta ohjelmoitujen tehtävien ulkopuolelle. (Roff, 2014, 2013.) Toisaalta taas Sheridan ja Verplank määrittävät robottiautonomisuuden siten, että jos robotti kykenee tekemään kaikki tehtävänsä liittyvät toimenpiteet itsenäisesti, ilman että ihmisen tarvitsee tai pitää ottaa osaa sen toimintaan, niin silloin laite toimii autonomisesti. (Noorman & Johnson 2014, 57.) Tämän määritelmän mukaan MK – 15 Phalanx -järjestelmää voitaisiin pitää autonomisena. Ongelmana kuitenkin Noorman ja Johnson huomauttaa tähän, että ihmisen kontrolli pysyy kyseisien autonomisten järjestelmien luomisessa ja ohjelmoinnissa (Noorman & Johnson 2014, 57.) Täten vaikka Sheridanin ja Verplankin autonomisuuden skaalassa robotti olisi yläpäässä, niin ihmiset olisivat silti sen toiminnasta kontrollissa kehityksen kautta. Noorman ja Johnson (2014) esittävät robottiautonomiasta huomion, että robottien autonomia on oletus, jolla on voimaa vain silloin, kun niiden sisäinen toiminta pidetään pimennossa, niin kutsuttuna ”mustana laatikkona”. Heti kun laatikko avataan, ja nähdään, miten robotin autonomia on ihmisen rakentama, niin vastuukysymys muuttuu selkeästi. (Noorman & Johnson 2014, 60.)

Käyttäjän vastuu

Koska moraalisen vastuun metafyyssisiin kysymyksiin on mietitty ratkaisuja jo pidemmän aikaa, eikä mitään konklusiivista olla saatu kehitettyä, ja silti ihmiset syyttävät ja antavat kehuja toisiensa teoista, niin tästä syystä perustan sotarobotiikan käytön vastuutarkastelun P.F.Strawsonin vuonna 1963 kirjoittamaan *Freedom and Resentment* artikkeliin, jossa keskeisimpänä tekijänä hän näkee taipumuksemme reaktiivisille asenteille. Ihmisillä on tapana reagoida eri tavoin siihen, miten muut toimivat heitä kohtaan. Tässä kohtaan tarkastelu siirtyy aidon metafyyssisen vastuun selvittämisestä ihmistenvälisiin käytäntöihin ja konteksteihin, joissa kyseiset asenteet ilmenevät. Strawsonin

tarkastelu sivuuttaa vapaan tahdon ja keskittää moraalisen vastuun enemmän sosiaalisiin käytäntöihin, erityisesti syyttämisen ja kehumisen muodoissa (Nuzzi & Sio 2016, 171.) Strawson asettaa teoriansa ytimeen ihmisen itserakkauten (self-love), itsetunnon (self-esteem), tai turhamaisuuden (vanity), ja kuinka näitä pystyisi kunnioituksella hyväilemään, tai satuttamaan välinpitämättömyydellä. Toisaalta voitaisiin puhua rakkaudesta, tai sen puutteesta johtuvasta turvattomuuden tunteesta. Näihin liittyen Strawsonilla on ollut ideana osoittaa, kuinka paljon meille merkitsee se, että miten ihmiset kohtelevat meitä, ja millaisella reaktiivisella asenteella he meitä kohtelevat. Kohtelu voi aiheutua sekä välittävyydestä, hyväntahtoisuudesta tai kunnioituksesta, ja toisaalta pahantahtoisuudesta, välinpitämättömyydestä tai halveksinnasta. (Strawson 1963, 3.)

Yhteiskunnallisessa lähestymistavassa moraaliseen vastuukysymykseen on tuotava moraalinen viitekehys, mitä vasten syyttämisen ja kehumisen käytäntöjä tarkastellaan. Alex Leveringhaus ottaa artikkelissaan *Drones, Automated Targeting and Moral Responsibility* tällaiseksi viitekehykseksi Oikeutettu Sota -teorian. Oikeutettu Sota -teoriassa on kaksi säännöstelevää osiota, jotka rajoittavat aseellista voimankäyttöä, sekä yksi säännöstelevä osio, joka ottaa kantaa sodan jälkipuintiin. Nämä osiot ovat nimeltään *jus ad bellum* (oikeudenmukainen sodanjulistus), *jus in bello* (oikeudenmukainen sodankäynti) ja *jus post bellum* (oikeudenmukaisuus sodankäynnin jälkeen). (Nuzzi & Sio 2016, 171.)

Jus in bello eli oikeutetun sodankäynnin osiossa keskeisiä kriteerejä ovat väkivallan käytön kohteiden erottaminen (distinction), väkivallan käytön suhteellisuus ja aseellinen tarpeellisuus. Nämä säännökset antavat suuntaa sille, milloin voimme nähdä robottiaseiden käyttäjän olevan vastuussa epäeettisestä sodankäynnin käytännöstä. (Nuzzi & Sio 2016, 171.) Kohteiden erottamisella tarkoitetaan, että sodankäynnissä osapuolten sotilaiden on eroteltava vastapuolen sotilaat ja siviilit toisistaan. Siviilit ovat moraalisesti immuuneja kohtisuorille hyökkäyksille, ja ainoastaan epäsuorasti aiheutuneita siviilihenkiä menetyksiä voidaan pitää anteeksi annettavana. Voimakeinojen suhteellisuus taas viittaa siihen, että sodan osapuolet eivät saa käyttää ylimitoitettuja voimakeinoja tavoitteidensa saavuttamiseen. Täten eettisenä käytäntönä ei pidetä esimerkiksi sokkopommittamista tai ylipommittamista, joista eniten on joutunut usein kärsimään juuri siviilijoukot. Aseellisella tarpeellisuudella taas tarkoitetaan, että sotaa käyvät eivät saa aiheuttaa tarpeetonta vahinkoa. (Tzafestas 2016: Nuzzi & Sio 2016, 171.) Muita Oikeutettu Sota -teorian säännöstöjä ovat muun muassa kontrolloidut aseet, jonka mukaan osapuolet eivät saa käyttää aseita, jotka ovat ”pahoja itsessään” kuten biologiset aseet, ja näiden aseiden pitää olla myös heidän hallittavissa. (Tzafestas 2016.) Oletettavaa tietenkin on, että sotilaiden käyttäessä robotteja sodankäynnin välineenä, ne on testattu toimimaan tietynlaisessa ympäristössä, jossa niiden halutaan suorittavan tehtävänsä. Täten

myös oikeutetun sodankäynnin säännöstö edellyttää, että robotit ovat testattu kattavasti vaadittua tehtävää varten.

Strawsonin sosiaalisiin käytäntöihin perustuvassa teoriassa tärkeää ei ole pelkästään se, millaisella asenteella tai intentiolla toinen kohtelee meitä. Keskeistä on myös se, milloin katsomme toisen olevan moraalisesti vastuullinen toimija. Ehtoja tälle ovat toimijan moraalinen kyvykkyys, moraalinen käsittämiskyky ja toiminnan kontrolli. Moraalinen kyvykkyys tarkoittaa, että toimijan pitää kyetä ymmärtämään mitä hän teki, ja miksi häntä pidetään vastuullisena toiminnastaan. Henkilö saattaa olla moraalisesti kyvytön esimerkiksi, jos hän ei ole vielä kehittynyt lapsesta aikuiseksi, jos hän on kieroutunut tai jos hän on muuten neuroottinen. Moraalisella käsittämiskyvyllä tarkoitetaan, että toimijan tulee ymmärtää moraalisesti tilanne, jossa hän toimii. Kontrollilla tarkoitetaan, että toimijalla täytyi olla mahdollisuus toimia toisin, jotta häntä voidaan pitää vastuullisena teoistaan. Nämä asiat voidaan katsoa määrittelevän moraalisesti vastuullisen toimijan, mutta haasteellista tietenkin on, että kuinka hyvin nämä asiat pystytään todentamaan tilanteessa kuin tilanteessa. (Nuzzi & Sio 2016, 171: Strawson 1963, 5) Kontrollin tapauksessa enää emme puhu siis metafysisestä valinnan mahdollisuudesta, vaan enemmän Humelaisesta valinnan mahdollisuudesta.

Ennen toisen maailmansodan jälkeen käytyjä Nürnbergin oikeudenkäyntejä sotilaiden täytyi vain osoittaa, että he noudattivat toimiessaan valtuutettuja käskyjä vapautuakseen moraalisesti vastuusta. Vuosina 1945–1946 käydyissä oikeudenkäynneissä kuitenkin lisättiin kaksi kriteeriä moraalisen vastuun määrittämiseen. Ensimmäinen kriteeri on moraalinen havaintokyky (moral perception), jonka mukaan toimiessaan sotilaan tulee pystyä esittämään, että hän joko kykeni tai ei kyennyt hankkimaan kaikkea moraalisesti relevanttia informaatiota arvioidessaan aseellista voimankäyttöään tietyissä tilanteissa. Toinen kriteeri on moraalinen vaihtoehto (moral choice), joka vaatii, että sotilaan on osoitettava, oliko hänellä vaihtoehtoa annetun käskyn tottelemisen suhteen. Enää ei siis riittänyt, että sotilas osoitti toimineensa esimiehensä käskyistä, jos hänellä olisi ollut mahdollista valita moraalisempi vaihtoehto. (Nuzzi & Sio 2016, 172: Lin et al. 2011, 152.)

Moraalinen havaintokyky edellyttää toimijalta, että hän tuntee jaetut moraaliset viitekehykset, joita sodankäynnissä noudatetaan. Näitä moraalisia arvoja hänen tulee osata peilata todellisuuteen ja löytää niiden avulla relevantit moraaliset tosiasiat ympäristöstään. Esimerkiksi *Jus in bello* -säännöksissä erottelu (discrimination) perustuu ajatukselle, että siviilit on pystyttävä erottelemaan sotilaista, ja heihin kohdistettu väkivalta on kiellettyä. Moraalisesti kyvykäs toimija kykenee ymmärtämään tämän tosiasian. Moraalisesti kyvykkään toimijan pitää kyetä myös hankkimaan kaikki saatavilla oleva tarpeellinen tieto, jotta hän pystyy arvioimaan kyseisen tilanteen missä hän toimii. Kyseistä informaatiota toimijan on kyettävä lopulta vertailemaan tuntemaansa moraaliseen koodistoon, ja

valittava siitä päätöksenteolle oleellinen osa. Jos sotilas tietää tietyssä tilanteessa kaikki tosiasiat ja pystyy tunnistamaan niistä moraalisesti olennaiset, hänen voidaan katsoa olevan sekä kyvykäs että tietoinen (aware) moraalisesti. Jos esimerkiksi sotilas on tilanteessa, jossa henkilö ei kohdistu häntä kohtaan epäoikeudenmukaista uhkaa, eikä hänen voida katsoa myöskään olevan vastuussa tällaisen uhan luomisesta, nähdään hänen olevan viaton ja täten myös moraalisesti immuuni hyökkäyksiä kohtaan. (Nuzzi & Sio 2016, 172.)

Tässä kohtaa kysymykseksi asettuu, voiko robotiikka vaikuttaa sotilaiden kykyyn tehdä moraalisia päätöksiä. Esimerkiksi etäältä ohjattujen hyökkäyslentokoneiden ohjaajista on spekuloitu, että välimatka sotarintamalle vaikuttaa muun muassa tappamisen vastustamisen heikentymiseen (Sharkey 2010, 3). Psykologit ja osa armeija-analyttikoista väittävät, että etäältä tappaminen ja hengenvaaralta välttyminen saattaisi johtaa tunteellisesti ja moraalisesti erkaantumiseen, jolloin sotilaat eivät kykenisi enää näkemään tekojensa vakavuutta, joka taas puolestaan saattaisi johtaa tappamisen estojen keventymiseen (Krishnan 2009, 128). Robotiikka voi tuoda omana moraalisen haasteenaan myös täysin automatisoidun tai osaksi automatisoidun kohteen tunnistamisen (Nuzzi & Sio, 2016, 173).

Robotiikka saattaisi vaikuttaa saatavilla olevaan moraalisesti relevanttiin informaation, josta sotilas tekee moraalisesti tietoiset päätöksensä. Esimerkiksi sotalennokit nauhoittavat suuria määriä informaatiota, josta lennokit sitten suodattavat sotilaille ohjelmointinsa mukaisesti tietyn osan, jotta lennokkia operoiva ihminen pystyy käsittelemään sitä. Tämä kuitenkin saattaa vaikuttaa niin, että jokin tärkeä osa informaatiosta jää sotilalla saamatta. Toisaalta taas observoinnin automatisoiminen ei välttämättä ole pelkästään negatiivinen asia, sillä ihmiselle on haasteellista kyetä käsittelemään valtavat määrät informaatiota, jonka lennokka kykenee suodattamaan hänelle. Tämä saattaisi myös heikentää moraalisten tosiasioiden löytämistä. Niin pitkään, kun ihminen on mukana ampumisen päätöksenteossa, niin hän pystyy syrjäyttämään robotin päätöksen kohdistaa väkivaltaa jotakin kohdetta kohtaan. Tulevaisuudessa vaarana on myös se, että roboteista tulee teknologian ansiosta niin nopeita ettei ihmisen kognitiiviset kyvyt välttämättä enää riitä tilanteiden moraaliseen arviointiin. (Nuzzi & Sio 2016, 174-176.)

Moraalisen havaintokyvyn ei kuulu ulottua pelkästään tilanteen aikana tehtyyn arviointiin, vaan myös arvioidut tulevaisuuden riskit robotiikan käytöstä tulisi laskea siihen mukaan. Tämä on keskeistä esimerkiksi juuri silloin, kun robotiikasta tulee liian nopeaa ihmisen hallittavaksi sen käytössä ollessa (Nuzzi & Sio 2016, 176.) Tällöin sotilaskomentaja asettaa riskin sille, että jokin tietty lopputulema tapahtuu. Tätä Himmelreich kutsuu artikkelissaan *Responsibility for Killer Robots* probabilistiseksi eli todennäköiseksi tai mahdolliseksi lopputulemaksi. Voi olla, että komentajalla ei ole hallinnassaan

yksityiskohtaisesti mitä robotti saattaa ottaa kohteikseen, mutta hänellä on silti kontrolli siitä, että robotti otetaan käyttöön ja sille annetaan jokin yleinen tehtävä suoritettavaksi. Himmelreich valaisee tätä ajatusta antamalla esimerkin, jossa tulevaisuuden lennokkirobotti on asetettu partioimaan aluetta, jossa sen tehtävä on pommittaa legitiimejä kohteita. Robotti pommittaa ryhmän vihollisotilaita, jotka ovat päättäneet luovuttaa, mutta robotti ei tunnista heitä oikein. Tällöin syntyy niin sanottu vastuukuilu, jolloin ei ole selvää, että kuka on vastuussa, koska komentaja ei käskenyt pommittaa juuri tätä kohdetta, mutta koska komentaja on asettanut tietoisesti riskin robotin käyttöönoton yhteydessä, että näin voi käydä, hänen nähdään olevan vastuussa. (Himmelreich 2019, 2,5.)

Himmelreich antaa seuraavan määritelmän kontrollille, joka kuuluu näin:

Vahva (robust) jäljittävä kontrolli:

1. on olemassa käsky, jonka a voi antaa siten, että
2. jos a päättää antaa käskyn, niin x tapahtuu (kaikissa relevantisti samanlaisissa tilanteissa), ja
3. jos a ei antaisi kyseistä käskyä, niin x ei tapahtuisi (kaikissa relevantisti samanlaisissa tilanteissa).

Tämän määritelmän mukaan, jotta lopputulema x voitaisiin nähdä olevan toimijan a kontrollissa, ei riitä pelkästään, että käskyn antaminen nostaisi tapahtuman todennäköisyyttä, vaan myös jos käsky jää antamatta, niin lopputulemaa ei tapahdu. Lopputulemasta voi siis päätellä, annettiinko käskyä vai ei. Vahvaan jäljitettävään kontrolliin voidaan sisältää mahdollisuus, että käskyn antamalla toimija asettaa samalla riskin, että asiat voivat mennä pieleen. Lopputulema x voidaan ilmaista disjunkttiivisesti ”joko tapahtuma E_1 tapahtuu tai tapahtuma E_2 tapahtuu”, jossa E_1 ja E_2 ovat kuvauksia tietyistä tapahtumista luonnollisella kielellä ilmaistuna. (Himmelreich 2019, 6.) Tällöin lopputulema voi sisältää myös mahdollisuuden, että lennokkirobotti pommittaa epälegitiimejä kohteita. Himmelreichin mukaan esimerkiksi tapauksissa, joissa robottia on käsketty partioimaan aluetta, tiettyyn (particular) lopputulemaan ”tiettyä kohdetta pommitetaan” ei pystytä vaikuttamaan, mutta probabilistiseen ”jotakin kohdetta pommitetaan tai mitään kohdetta ei pommiteta” pystytään. Tämä probabilistinen lopputulema nähdään olevan komentajan kontrollissa. (Himmelreich 2019, 8.)

Sotakomentajan moraaliseen kyvykkyyteen vaikuttaa siis, miten hyvin ja kattavasti robotit ovat testattu ennen niiden käyttöönottoa, jotta hänellä on tiedossa mahdollisen virheen todennäköisyydet. Yhdysvaltojen puolustusvoimien vuoden 2011 etenemissuunnitelmassa on asetettu vaatimuksia, että robottien on oltava luotettavia ja ennustettavissa olevia, jotta niiden voitaisiin antaa operoida itsenäisesti. Etenemissuunnitelman mukaan laitteiden luotettavuuden tarkistukseen on kehitettävä verifikaatio- ja validointimenetelmiä, jota pidetään yhtenä vaikeimmista teknisistä esteistä kyseisten

laitteiden käyttöönotossa ja hyödyntämisessä. (Noorman & Johnson 2014, 59.) Himmelreich (2019) ottaa vastuun periaatteeksi myös ennustettavuuden (foreseeable), joka edellyttää, että komentaja pystyy näkemään ennalta, että autonominen asejärjestelmä saattaa pommittaa epälegitiimejä kohteita (Himmelreich 2019, 14). Tämä periaate on jonkin verran yhteensopiva oikeutettu sota -periaatteiden kanssa, sillä moraalinen havaintokyky edellyttää juuri sitä, että toimija kykenee keräämään kaiken moraalisesti relevantin informaation, jotta hän pystyy arvioimaan tilanteen, missä toimii. Yleisten armeijakäytäntöjen mukaan, komentajien annetaan tietää testien jälkeen, mikä virhetoiminnan todennäköisyys on (Himmelreich 2019, 14).

Probabilistiset ja koneoppimisalgoritmit, jotka mahdollistavat robotin mukautumisen ja oppimisen, eivät aiheuta sattumanvaraista toimintaa, vaan se miten ne pystyvät mukautumaan on kehittäjien ja käyttäjien rajoitettavissa. Tällöin vastuukysymysten täytyy keskittyä siihen, millaisia päätöksiä ja strategioita on käytetty laitteiden kehityksessä tai käytössä. Kehittäjät määrittävät tietokoneohjelmistollaan robotin toiminnan laajuuden, ja tästä syystä tärkeää olisi kehittää käytäntöjä sille, miten ja milloin vastuu määritetään näille henkilöille. Ihmis-robotti-kokonaisuuksille on olemassa eteenpäin ja taaksepäin katsovia vastuukäytäntöjä, joissa eteenpäin katsova vastuukäytäntö koskee sitä, miten eri tehtävistä ja velvollisuuksista jaetaan vastuu ihmis- ja konekomponenttien kesken systemissä. Taaksepäin katsovassa käytännössä vahingon tai virheen sattuessa tarkastellaan näitä ennalta asetettuja vastuualueita, ja selvitetään mistä tai kenestä virhe johtui. Taaksepäin jäljittämisen jälkeen mahdollista on tehdä myös päätös, että jokin tehtävä siirretään ihmistoimijan tai koneen vastuulle. (Noorman & Johnson 2014, 60.)

Pohdinta

Metafyysisellä tasolla tarkasteltuna moraalinen vastuu on edelleen siis mysteeri, sillä vaikuttaisi siltä, että determinismillä on vahva peruste sen oikeellisuudelle, jolloin kenenkään ei voitaisi nähdä olevan aidosti vastuussa toiminnastaan. Jos käytännön tasolla kuitenkin tahdotaan, että sotarobottien käytössä saadaan selväksi, kuka on vastuussa niiden toiminnasta, niin tällöin kontrollin ja autonomian käsitteet ovat keskeisiä. Jos robotin nähdään olevan moraalisesti vastuussa toiminnastaan, niin silloin sen pitäisi omata autonomia, eli kyky toimia itsenäisesti. Vaikka robottien nähdään toimivan enemmän itsenäisesti, niin samalla niiden toimintaa edelleen ohjaavat ihmiset ohjelmoinnin, testauksen ja käyttötapojen kautta.

Hallitsemattomissa olevien robottien kehittäminen olisi edelleen ihmisten päätettävissä, ja tällaiseen lopputulokseen ei todennäköisesti päädytä, koska armeijat tavoittelevat laitteistoja, joiden voidaan nähdä olevan ennustettavia ja luotettavia. Nykyhetkessä on vaikea kuvitella, miten roboteilla saatettaisiin nähdä olevan vapaa tahto, mutta jos roboteista tulee joskus ihmisenkaltaisia, niin tämä saattaa hämärtää ihmisten intuitioita aiheesta.

Oikeutettu sota -teorian ja P.F. Strawsonin reaktiiviset asenteet -teorian kautta pystymme tuomaan moraalisen viitekehyksen sille, mitä sääntöjä ja edellytyksiä sotilaille on sodankäynnissä. Robotiikka voi vaikuttaa sotilaiden moraaliseen havaintokykyyn operoidessa ja suodattaessa sotilaille tietoja esimerkiksi oikeiden kohteiden tunnistamista varten. Tämä voi kuitenkin olla joko positiivinen tai negatiivinen, riippuen siitä kuinka hyvin robotti aidosti kykenee auttamaan kohteiden tunnistamisessa. Sotilaiden moraaliseen kyvykkyyteen spekuloidaan vaikuttavan robotiikan mahdollistamat etäiskut, jolloin sotilas ei koe samoja negatiivisia vaikutuksia, mitä ihmisen tappamisesta usein aiheutuu lähitaisteluissa. Tämä saattaa johtaa siihen, että sotilaat tappavat helpommin kuin koskaan ennen. Tästä syystä tärkeää on tutkia, miten robotiikka vaikuttaa käyttäjänsä, jotta käyttäjät pystyvät tiedostamaan nämä vaikutukset.

Sotilaskomentajan moraalinen havaintokyky pitää ulottua sekä nykyhetkeen, että myös tulevaisuuteen robotiikan käytön riskien arvioimisessa. Kasvavan autonomisuutensa ja nopeutensa ansiosta robotit saattavat tehdä tekoja, joihin ihminen ei ole antamassa suoraa käskyä. Tällöin käyttäjän vastuulla on selvittää, mikä todennäköisyys on, että robottia käytettäessä se tekee jotakin, mitä käyttäjä ei toivo. Käyttäjä asettaa riskin, että robotti saattaa tehdä jotakin, esimerkiksi tunnistaa kohteensa, väärin. Ihmisten ja koneiden muodostamassa systeemissä voidaan käyttää myös eteenpäin

ja taaksepäin katsovia vastuukäytäntöjä, jolloin myös robottien kehittäjien vastuuta voidaan selvittää, ja tähän on tarve myös kehittää selkeitä vastuukäytäntöjä.

Lähteet

- Asaro, Peter. 2008. *How just could a robot war be?* Current Issues in computing and philosophy (pp. 50–64). Amsterdam, The Netherlands: IOS Press
- Ashrafian, Hutan. 2015. *Artificial Intelligence and Robot Responsibilities: Innovating Beyond Rights*, Science and Engineering Ethics, 317-326.
- Griffith, Meghan. 2013. *Free Will: the Basics*, Taylor & Francis Group.
- Himmelreich, Johannes. 2019. *Responsibility for Killer Robots*. Ethical Theory and Moral Practice 22: 731-747. <https://doi-org.libproxy.tuni.fi/10.1007/s10677-019-10007-9> (Luettu Toukokuu 3, 2020).
- Krishnan, Armin. 2009. *Killer Robots : Legality and Ethicality of Autonomous Weapons*, Taylor & Francis Group.
- Noorman, Merel & Johnson, Deborah G.. 2014. *Negotiating Autonomy and Responsibility in Military Robots*. Ethics and Information Technology, 16(1), 51-62.
<http://dx.doi.org.libproxy.tuni.fi/10.1007/s10676-013-9335-0> (Luettu Huhtikuu 29, 2020).
- McFee, Graham. 2000. *Free Will*, Taylor & Francis Group.
- Nucci, E. D., & Sio, F. S. D.. 2016. *Drones and responsibility : Legal, philosophical and socio-technical perspectives on remotely controlled weapons*. Taylor & Francis Group.
- Patrick Lin, James J. Hughes, Keith Abney, George A. Bekey, Colin Allen, Anthony Beavers, Paul Bello, Jason Borenstein, Selmer Bringsjord, and Marcello Guarini. 2011. *Robot Ethics: The Ethical and Social Implications of Robotics*, MIT Press. ProQuest Ebook Central, <https://ebookcentral.proquest.com/lib/tampere/detail.action?docID=3339351>.
- Roff, Heather. 2014. *The Strategic Robot Problem: Lethal Autonomous Weapons in War*. Journal of Military Ethics. 13(3):211-227. doi:10.1080/15027570.2014.975010
- Sharkey N. 2010. *Saying "No!" to Lethal Autonomous Targeting*. Journal of Military Ethics, 9(4), 369-383.
- Smilansky, S. 2011. *Hard Determinism and Punishment: A Practical "Reductio"*. Law and Philosophy, 30(3), 353-367. Retrieved August 20, 2020, from www.jstor.org/stable/41486987
- Sparrow, Robert. 2007. *Killer Robots*. Journal of Applied Philosophy, 24(1), 62-77.
<https://doi-org.libproxy.tuni.fi/10.1111/j.1468-5930.2007.00346.x> (Luettu Huhtikuu 27, 2020).
- Strawson, Galen. 2002. The Bounds of Freedom. *The Oxford Handbook on Free Will*, 441-460.
- Talbert, Matthew. 2016. *Moral Responsibility: An Introduction*, Polity Press.

Talbert, Matthew. *Moral Responsibility*, The Stanford Encyclopedia of Philosophy (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2019/entries/moral-responsibility/>>.

Tzafestas S.G.. 2016. War Roboethics. In: Roboethics. Intelligent Systems, Control and Automation: Science and Engineering, vol 79. Springer, Cham. https://doi-org.libproxy.tuni.fi/10.1007/978-3-319-21714-7_9

Waller, Bruce. 2014. *The Stubborn System of Moral Responsibility*. Cambridge, Massachusetts: The MIT Press