

Melisa Kütük

PERSOONALLISUUSPIIRTEET JA LUOTTAMUS ROBOTTEIHIN JA TEKOÄLYYN

Yhteiskuntatieteiden tiedekunta
Pro gradu -tutkielma
Maaliskuu 2021

TIIVISTELMÄ

Melisa Kütük: Persoonallisuuspiirteet ja luottamus robotteihin ja tekoölyyn
Pro gradu -tutkielma
Tampereen yliopisto
Yhteiskuntatutkimuksen tutkinto-ohjelma
Sosiaalipsykologia
Maaliskuu 2021

Robottien ja tekoölyn käyttö on lisääntynyt merkittävästi viimeisten vuosikymmenten aikana eri aloilla. Niitä käytetään yhä enemmän tilanteissa, jotka ovat vaarallisia tai jostain muusta syystä ihmisten ulottumattomissa. Erilaiset pelastustehtävät, lääketieteelliset operaatiot, puolustusvoimat tai avaruuden tutkimus ovat vain pieni osa siitä, mihin kaikkien robottien ja tekoölyn käyttö mahdollistaa. Tilanteissa, joissa ihmiset joutuvat tekemään yhteistyötä robottien kanssa, on luottamus keskeistä ja sekä liian korkea että liian matala luottamus voivat olla kohtalokkaita. Yksi luottamukseen vaikuttavista tekijöistä on ihmisten persoonallisuuspiirteet. Vaikka niiden merkitys luottamukseen vaikuttaa olevan kiistaton, yksittäisten piirteiden vaikutuksesta on aiemmissa tutkimuksissa saatu ristiriitaisia tuloksia.

Tämän tutkimuksen tavoitteena oli tutkia persoonallisuuspiirteiden yhteyttä robotteja sekä tekoölyä kohtaan tunnettuun luottamukseen. Samalla tutkittiin iän, sukupuolen, työn ja koulutuksen yhteyttä robotteja ja tekoölyä kohtaan tunnettuun luottamukseen sekä sitä, miten luottamus robotteihin eroaa luottamuksesta tekoölyyn. Tutkimus toteutettiin osana Tampereen yliopiston monitieteistä *Robotit ja me: vuorovaikutuksen fysiologinen, psykologinen ja sosiaalinen ulottuvuus* -hanketta. Luottamusta tutkittiin luottamuspelin avulla ja yhdysvaltalaisista koostuva aineisto ($n = 969$) kerättiin verkkokyselynä. Analyysiin käytettiin kuvailevia analyysijä ja lineaarista regressioanalyysiä.

Tekoölyryhmässä avoimuuden ja iän havaittiin ennustavan positiivisesti ja tunnollisuuden negatiivisesti luottamusta. Robottiryhmässä ainoastaan avoimuus oli tilastollisesti merkitsevästi yhteydessä luottamuksen kanssa. Havaitut yhteydet eivät olleet voimakkaita, mutta ne olivat tilastollisesti merkitseviä. Osallistujien luottamus ei vaihdellut robottien ja tekoölyn välillä, ja molemmille annettiin suunnilleen yhtä paljon rahaa. Tässä tutkimuksessa saadut tulokset vastaavat osittain aiemmin tehtyjä tutkimuksia. Kirjallisuudessa tulokset ovat kuitenkin paikoin hyvinkin ristiriitaisia ja lisää tutkimuksia tarvitaan etenkin persoonallisuuspiirteiden ja luottamuksen välillä, jotta saataisiin kehitettyä luotettavampia robotteja ja tekoölyä, jotka olisi helpompi hyväksyä.

Avainsanat: Luottamus, persoonallisuus, robotit, tekoöly

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

Sisällysluettelo

1 Johdanto.....	1
2 Luottamus	4
2.1 Luottamuksen määrittely	4
2.2 Luottamuksen vaikutukset robottien sekä tekoälyn välisessä vuorovaikutuksessa	7
2.3 Luottamuksen mittaaminen.....	10
2.4 Luottamukseen vaikuttavat tekijät	12
3 Persoonallisuuspiirteet.....	15
3.1 Persoonallisuuden määrittely	15
3.2 Persoonallisuuspiirteiden ja luottamuksen välistä yhteyttä tutkivat aiemmat tutkimukset.....	20
4 Tutkimuskysymykset ja hypoteesit	24
5 Aineisto ja menetelmät	25
5.1 Aineisto	25
5.2 Tutkimusetiikka	25
5.3 Muuttujat.....	26
5.3.1 Luottamus.....	26
5.3.2 Persoonallisuus.....	26
5.3.3 Taustamuuttujat.....	27
5.4 Analyysimenetelmät	29
5.4.1 Kuvailevat analyysit.....	29
Riippumattomien otosten t-testi	29
Korrelaatio.....	30
5.4.2 Lineaarinen regressioanalyysi.....	30
Lineaarisen regressioanalyysin ehdot	32
6 Tulokset	33
6.1 Kuvailevat tulokset	33
6.2 Regressioanalyysin tulokset.....	36
7 Pohdinta.....	38
7.1 Pohdinta	38
7.2 Rajoitteet ja jatkotutkimus	42
7.3 Loppupäätökset.....	43
Lähdeluettelo	45
Liitteet.....	55

Taulukot ja kuvat

Taulukko 1: Big Five -mallin ylä- ja alapiirteet.....	17
Taulukko 2: Kategoristen muuttujien kuvailevat tunnusluvut robotti- sekä tekoälyryhmissä	28
Taulukko 3: Jatkuvien muuttujien kuvailevat tunnusluvut.....	29
Taulukko 4: Muuttujien väliset korrelaatiot robottiryhmässä.....	34
Taulukko 5: Muuttujien väliset korrelaatiot tekoälyryhmässä.....	35
Taulukko 6: Lineaarisen regressioanalyysin tulokset koskien persoonallisuuspiirteitä ja taustamuuttujia sekä robotille/tekoälylle annettua rahamäärää	37
Kuvio 1: Luottamusmalli	7
Kuvio 2: Kalibraation, resoluution sekä automaation kyvykkyyden välinen yhteys automaatiota kohtaan tunnetun luottamuksen synnyssä	9
Liitteet	
Liite A: Regressiomallien vif-arvot robottiryhmässä	55
Liite B: Regressiomallien vif-arvot tekoälyryhmässä.....	55

1 Johdanto

Robottien ja tekoälyn käyttö on lisääntynyt merkittävästi viimeisten vuosikymmenten aikana eri aloilla ja teknologian kehittyessä niiden keskeinen asema näyttää tulevaisuudessa vain vahvistuvan (Makridakis, 2017). Populaarikulttuuri on pitkälti muokannut käsitteitämme ja ajatuksiamme roboteista sekä tekoälystä, ja usein luottamuksen vastakohtana on pelko liian älykkäiden robottien kapinasta (Coeckelbergh, 2011). Todellisuudessa robotit ja tekoäly mahdollistavat ihmisten saavuttamattomissa olevia asioita, tarjoten laajan potentiaalin eri aloille. Niiden käyttö näkyy muun muassa erilaisissa pelastusoperaatioissa, maanpuolustuksessa (Chen ym., 2010) sekä lääketieteellisissä leikkauksissa (Tsui & Yanco, 2007). Robotteja käytetäänkin yhä enemmän tilanteissa, jotka ovat vaarallisia tai jostain muusta syystä ihmisten ulottumattomissa (Hancock ym., 2011). Erilaiset pelastustehtävät, lääketieteelliset operaatiot, puolustusvoimat tai avaruuden tutkimus ovat vain pieni osa siitä, mihin kaikkeen robottien ja tekoälyn käyttö mahdollistaa. Näissä tilanteissa, joissa ihmisten pitää toimia yhteistyössä robottien kanssa, on luottamus keskeistä (Billings ym., 2012) ja sekä liian korkea että liian matala luottamus voivat tällaisissa tilanteissa olla kohtalokkaita (Merrit ym., 2012).

Luottamuksella on tärkeä sija vuorovaikutuksessa ja esimerkiksi yhteiskunta ja sen instituutiot, kulttuurit sekä hallitukset pohjautuvat kaikki osittain luottamukseen (Hoff & Bashir, 2015). Vaikka ihmistenvälisen sekä ihmisten ja robottien tai tekoälyn välisen luottamuksen välillä on huomattu samankaltaisuuksia, poikkeavat ne myös monelta osin toisistaan (Lee & See, 2004). Mitä riippuvaisemmiksi robotiikasta sekä tekoälystä tulemme, sitä ajankohtaisemmaksi luottamuksen tutkimisestakin tulee. Ihmisten ja robottien välinen yhteistyö ja siihen perustuvat työryhmät, joissa molemmat ovat mukana, ovat yleistyneet ja niiden toimivuuden takaaminen perustuu pitkälti luottamukseen. Tiimin sisäinen vuorovaikutus ja jäsenten luottamus ovat avainasemassa tavoitteiden ja määränpään saavuttamisessa (Billings ym., 2012). Luottamuksen merkitys korostuu myös teknologian hyväksymisessä ja niihin turvautumisessa (Lewandowsky ym., 2000). Tutkimusten mukaan ihmiset ovat vastaanottavaisempia esimerkiksi robotteja kohtaan, kun heillä on aiempaa kokemusta niiden kanssa (Bartneck ym., 2007). Tarkempi ymmärrys luottamuksen

muodostumiseen liittyvistä inhimillisistä tekijöistä voi auttaa esimerkiksi ennustamaan tai kontrolloimaan ihmisten turvautumista roboteihin tai tekoölyyn.

Robotti on käsitteenä hyvin laaja ja sitä on mahdollista määritellä hyvin monin tavoin. Samoin ihmisillä on erilaisia ja hyvinkin vaihtelevia käsityksiä puhuttaessa roboteista tai tekoölystä. Kuitenkin se useimmiten määritellään fyysisen olemassaolon sekä älyllisen toimijuuden mukaan (Richards & Smart, 2013). Näin ollen se eroaa tekoölystä konkreettisten piirteiden kautta. Robotin pitäisi myös pystyä havaitsemaan ympäristöönsä ja tehdä sen mukaan päätöksiä seuraavasta toiminnosta. Robotiikka on lisäksi mahdollista jakaa eri kategorioihin sen käyttötarkoituksen mukaan, kuten industriaaliseen robotiikkaan (esim. automaatio autoalalla), ammattikäyttöön tarkoitettuun robotiikkaan (esim. hylättyjen miinojen etsiminen) sekä henkilökohtaiseen robotiikkaan (esim. hoivatyö) (ISO 8373, 2012).

Myös tekoöly on määritelty monin eri tavoin ja muuttuu alati teknologian kehittymisen myötä. Tärkein käsite, jota tekoölylle on liitetty, vaikuttaa kuitenkin olevan tavoitteiden saavuttaminen erilaisissa, monimutkaisissa ympäristöissä, maksimoiden onnistumisen todennäköisyyttä (Goertzel, 2006; Legg & Hutter, 2006). Tekoölyllä on myös kognitiivisia taitoja, kuten taito oppia, sopeutua ja kehittyä (Voss, 2005; Wang, 1995). Kolmas paljon käytetty käsite ihmisten ja teknologian välisen luottamuksen tutkimuksessa on ollut automaatio, joka on määritelty ”teknologiaksi, joka aktiivisesti valitsee dataa, muuntaa tietoa, tekee päätöksiä sekä kontrolloi prosesseja” (Lee & See, 2004, s. 50). Näin ollen esimerkiksi tekoöly on usein myös osa automaatioprosessia, vaikkei aina näin olekaan. Luottamuksen tutkimuksessa luottamusta automaatioon on yleisesti tutkittu enemmän kuin pelkästään roboteihin tai tekoölyyn, mutta tämä tutkimustieto tarjoaa hyvän pohjan ymmärtää luottamusta ja mahdollistaa tiedon soveltamisen muihin ulkopuolisiin toimijoihin, kuten roboteihin tai tekoölyyn (Hancock ym., 2011).

Luottamukseen vaikuttavat monet sisäiset (esimerkiksi persoonallisuus) sekä ulkoiset (esimerkiksi kulttuuriset erot) tekijät (Mayer ym., 1995; Moray ym., 2000). Vaikka persoonallisuuspiirteiden on huomattu vaikuttavan luottamuksen kehitykseen (Schaefer ym., 2016), ihmisiin liittyvien tekijöiden ja näistä erityisesti persoonallisuuspiirteiden tutkimus on ollut pirstaleista ja tulokset ovat olleet ristiriitaisia (Robert ym., 2020). Persoonallisuuspiirteiden todellisesta vaikutuksesta ei näin ollen ole olemassa yhtenäisiä

tuloksia. Sisäisten tai yksilötekijöiden vaikutus luottamukseen on keskeistä, koska ne saattavat vaikuttaa robotteihin ja tekoölyyn turvautumiseen sekä riippuvaisuuteen tavoilla, jotka eivät ole suoraan tekemisissä robottien tai tekoölyn omien ominaisuuksien kanssa (Lee & See, 2004). On tärkeää ymmärtää persoonallisuuden merkitys luottamuksen muodostumisessa, koska se auttaa ymmärtämään ihmisten suhtautumista uutta teknologiaa kohtaan ja voi tarjota mahdollisuuden kehittää robotteja ja tekoölyä helpommin hyväksyttäväksi. Tutkimuksista saadun tiedon valossa on helpompi vastata kohdattuihin haasteisiin ja ongelmiin ja kehittää teknologiaa esimerkiksi yhteiskunnallisesti hyödyllisempään suuntaan.

Tämä tutkimus toteutettiin osana Tampereen yliopiston monitieteellistä *Robotit ja me: vuorovaikutuksen fysiologinen, psykologinen ja sosiaalinen ulottuvuus* -hanketta (2018–2020). Tutkimuksen tavoitteena on selvittää, miten persoonallisuuspiirteet ovat yhteydessä robotteja ja tekoölyä kohtaan tunnettuun luottamukseen sekä tarkastella luottamukseen yhteydessä olevia muita taustatekijöitä. Samalla tarkoituksena on tutkia ja verrata robotteja ja tekoölyä kohtaan tunnetun luottamuksen eroja. Tässä tutkielmassa käydään aluksi läpi luottamuksen ja persoonallisuuden teoreettista pohjaa ja aiempia tutkimuksia, jonka jälkeen esitellään aineisto sekä käytetyt menetelmät. Tämän jälkeen raportoidaan analyyseistä saadut tulokset, joita pohditaan viimeisessä osiossa teorian ja aiempien tutkimusten valossa. Lopuksi käydään läpi tutkimuksen rajoitteet, suositellaan jatkotutkimuksia sekä esitellään johtopäätökset.

2 Luottamus

2.1 Luottamuksen määrittely

Luottamus on yksi vuorovaikutuksen kulmakivistä, jota on tutkittu laajasti ja joka on kiinnostanut eri tutkimusaloja jo pitkään. Tästä huolimatta sen määritelmästä ei ole tutkijoiden kesken päästy yksimielisyyteen. Luottamusta on usein tarkasteltu asenteen, aikomuksen tai käyttäytymisen näkökulmasta. Lee ja See (2004) määrittelevät luottamuksen ”asenteena, joka auttaa henkilöä saavuttamaan tavoitteensa epävarmassa ja haavoittuvassa tilanteessa” ja heidän mukaansa tätä määritelmää voi käyttää sekä ihmistenvälisessä että automaatiota kohtaan tunnetussa luottamuksessa. Myös tulokset tai luottamuksen kohteena olevan toimijan käyttäytymistä koskevat odotukset nousevat usein esiin luottamuksen määrittelyssä (Lee & See, 2004). Mayer, Davis ja Schoorman (1995) määrittelevät ihmisten välisessä vuorovaikutuksessa luottamuksen valmiutena olla haavoittuvainen toisen henkilön teoille ja tämä onkin kirjallisuudessa yksi käytetyimmistä määritelmistä (Rousseau ym., 1998). Haavoittuvuuden läsnäolo merkitsee, että henkilöllä on jotain menetettävää ja näin ollen luottamukseen sisältyy aina jonkinlainen riski. Riski onkin olennainen käsite luottamuksen tutkimuksessa sekä sen määrittelyssä. Deutschin (1958) mukaan luottamus ja riskinotto ovat toistensa vastakohtia ja riskejä otetaan vain, jos positiivisen tuloksen saamisen mahdollisuus on korkeampi kuin negatiivisen tuloksen saamisen. Riskinotto on samalla luottamuksen edellytys ja luottamuksen tarve esiintyy vain tilanteissa, joihin liittyy tietty riski. Moni tutkija määrittelee luottamuksen riskin avulla, koska se on läsnä kaikissa luottamukseen liittyvissä tilanteissa (Mayer ym., 1995).

Luottamusta on tutkittu ja huomioitu eri tieteenalojen piireissä jo pitkään. Jokainen näistä aloista, kuten psykologia, antropologia ja poliittinen historia lähestyvät luottamusta eri näkökulmista (Lewicki & Bunker, 1996). Worchelin (1979) mukaan eriävät lähestymistavat on kuitenkin mahdollista jakaa vähintään kolmeen ryhmään: persoonallisuustutkijoiden, sosiologien ja ekonomistien sekä sosiaalipsykologien näkökulmaan. Persoonallisuustutkijat tarkastelevat persoonallisuuspiirteiden vaikutusta luottamukseen ja luottamus mielletään uskona ja odotuksena, joka on osa ihmisen persoonallisuutta. Persoonallisuustutkijoiden mukaan luottamus on yksi, bipolaarinen konstruktio eli luottamus ja epäluottamus mielletään

erillisinä ja vastakkaisina asioina. Sosiologit ja ekonomistit taas pitävät luottamusta institutionaalisenä ilmiönä, jonka mukaan tulevaisuuden vuorovaikutus perustuu suoriin sekä epäsuoriin sääntöihin ja normeihin, kun taas sosiaalipsykologit keskittyvät luottamukseen ihmisten- ja ryhmienvälisenä ilmiönä (Lewicki ym., 1998).

Myös sosiaalipsykologian saralla luottamus ja riski nähdään vastakkaisina termeinä sosiaalisissa suhteissa. Luottamukseen liittyy turvallisuuden tunne ja usko siihen, että tarjottu tieto on todenmukaista (Bierhoff & Vornefeld, 2004). Suomalaiset sosiaalipsykologit Seppälä, Olakivi ja Pirttilä-Backman (2012) jäsentelevät luottamuksen ja sen tutkimuksen neljään eri tasoon käyttäen Willem Doisen (1986) sosiaalipsykologisen selittämisen tasoja. Nämä luokittelut jakautuvat yksilönsisäiseen, yksilöiden väliseen, ryhmien väliseen sekä ideologiseen tasoon. Yksilönsisäinen luottamus sisältää esimerkiksi älykkyyden, persoonallisuuden sekä mielialan. Yksilöiden sisäisessä luottamuksessa korostetaan esimerkiksi vastapuolen luotettavuuden käsitettä ja sen relationaalisuutta; luottamus on olemassa aina suhteessa johonkin. Luottamuksen ilmeneminen ryhmien sisällä ja välissä taas sisältää ryhmäsamaistumisen, jolloin omia ryhmäläisiä pidetään muita luotettavampina. Ideologisessa ja kulttuurisessa tasossa keskitytään kollektiivisiin ja individualistisiin kulttuureihin sekä luottamuksen asemaan poliittisena ja moraalisenä kohteena. Nämä kaikki ovat siis tärkeitä tekijöitä luottamuksen synnyssä ja sen muodostumisessa (Doise 1986; Seppälä ym., 2012).

Luottamus on mahdollista jakaa myös yksilökeskeisiin sekä sosiaalisiin teorioihin riippuen siitä, liitetäänkö luottamukseen vaikuttavat tekijät enemmän yksilöön vai sosiaalisiin suhteisiin ja yhteiskuntaan. Tämä käsitys on lähtöisin 1950–1960-luvun yhdysvaltalaisen sosiaalipsykologian koulukunnasta, johon esimerkiksi Erikson, Allport, Cattell ja Rosenberg vaikuttivat vahvasti. Luottamus nähdään pysyvänä piirteenä, joka on omaksuttu lapsena ja joka muuttuu hitaasti vain kokemusten ja erityisesti traumaattisten tapahtumien kautta. Sosiaalinen luottamus nähdään osana laajempaa persoonallisuutta, johon liittyy vahvasti esimerkiksi optimismi, elämäntyydytys ja usko yhteistyöhön. Luottamus on tämän näkökulman mukaan siis enemmän kytköksissä henkilön subjektiivisiin tunteisiin ja persoonallisuuteen, kuin ulkopuolisiin tekijöihin (Delhey & Newton, 2003). Delhey ja Newton (2003) kutsuvat näitä näkemyksiä persoonallisuusteoriaksi, jossa sosiaalinen

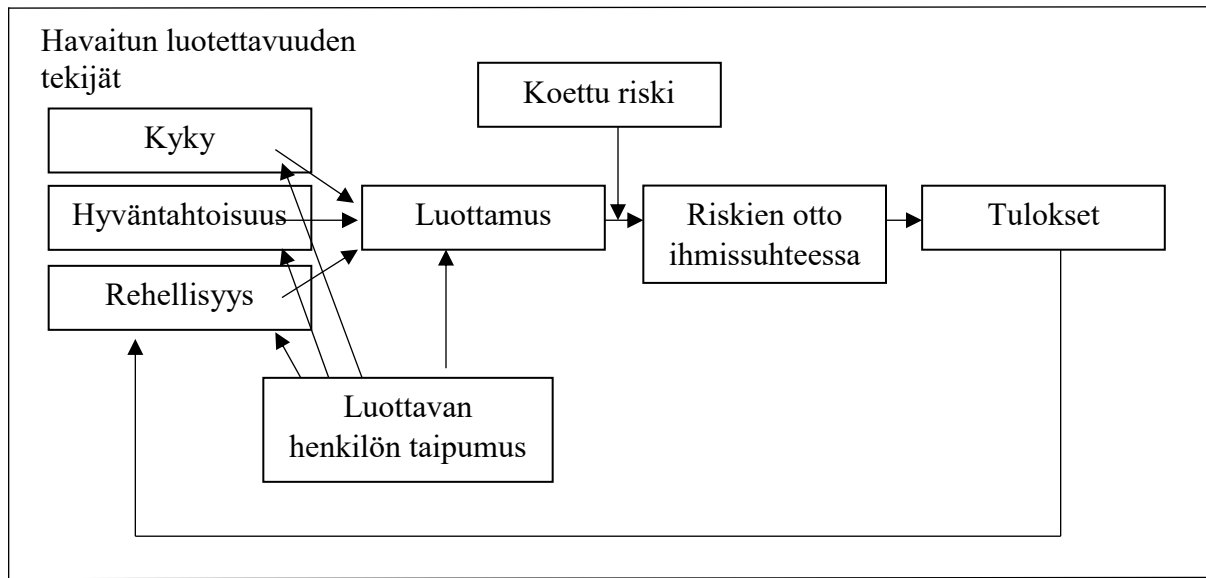
luottamus on yhteydessä muihin persoonallisuuden piirteisiin ja erityisesti optimismiin ja elämänhallintaan.

Vaikka luottamus on määritelty eri tavoin, liittyy siihen melkein aina kolme tekijää: luottaja, eli henkilö, joka luottaa, luotettava, eli luottamuksen kohde, sekä asia, joka on panoksena tai vaakalaudalla. Lisäksi tilanteeseen liittyy yleensä myös kannustin, jonka takia tehtävä suoritetaan (Hoff & Bashir, 2015). Päätös luottamuksen muodostamisesta sisältää myös kolme osaa. Ensinnäkin osapuoli suhteuttaa luottamuksen hyödyn luottamuksen toteutuessa sen haittoihin, jos luottamus ei toteudu. Toiseksi siihen liittyy oletus, että vastapuoli tekee jonkin asian, ja kolmanneksi, tapauksissa, joissa ei olla täysin varmoja, että vastapuoli tekee jonkin asian, osapuolen riskinotto ja riski joutua käytetyksi tulee kyseeseen (Ermisch ym., 2009).

Yksi laajasti käytetyistä luottamuksen teorioista on Mayerin ja kumppaneiden (1995) kattavan kirjallisuusanalyysin seurauksena kehittämä luottamusmalli. Tässä mallissa luottamusta ei nähdä riskin ottamisena, vaan enemmänkin haluna ottaa kyseinen riski. Mallin mukaiset luottamukseen vaikuttavat tekijät sekä luottajan että luotettavan kannalta on esitelty kuviossa 1. Mayer ja kumppanit (1995) löysivät siis kolme perustaa luottamukselle: kyky, hyväntahtoisuus ja rehellisyys, joista vähintään yhteen henkilön luotettavuus perustuu. Kyky on henkilön pätevyys ja taito tietyllä alalla, jonka takia henkilöön luotetaan. Luottamus on kuitenkin alasidonnaista, eikä välttämättä tarkoita, että kyseinen henkilö olisi luotettava muissa osa-alueissa. Hyväntahtoisuus kuvataan hyvien tekojen tekemisenä ilman omaa voitontavoittelua, jossa luotettava henkilö kokee luottavaa henkilöä kohtaan usein erityistä kiintymystä. Rehellisyydessä taas luottava olettaa, että luotettava noudattaa luottavan hyväksymiä, tiettyjä periaatteita.

Yksi tärkeä käsite luottamuksen tutkimuksessa on taipumus luottaa (engl. propensity to trust), jonka voi määritellä suhteellisen pysyvänä haluna ja taipumuksena luottaa muihin (Mayer ym., 1995). Taipumus luottaa muovautuu ja kehittyy ajan mittaan, kun henkilö on vuorovaikutuksessa muiden kanssa, saa vaikutteita yhteiskunnasta ja kuule muiden kokemuksista, alkaen muodostamaan uskomuksia koskien muita henkilöitä (Van Lange, 2015). Persoonallisuuspiirteistä sovinnollisuuden, neuroottisuuden ja avoimuuden on huomattu olevan yhteydessä taipumukseen luottaa (Alarcon ym., 2016). Tämän taipumuksen

voi soveltaa myös automatisaatiota, robotteja sekä tekoälyä kohtaan tunnetun luottamuksen tutkimukseen. Vaikka taipumus luottaa ja luottamus ovat lähekkäisiä käsitteitä, on kuitenkin syytä huomioida, etteivät ne ole sama asia ja tutkimuksissa niitä onkin käsitelty erillisinä käsitteinä.



Kuvio 1

Luottamusmalli (Mayer ym., 1995).

2.2 Luottamuksen vaikutukset robottien sekä tekoälyn välisessä vuorovaikutuksessa

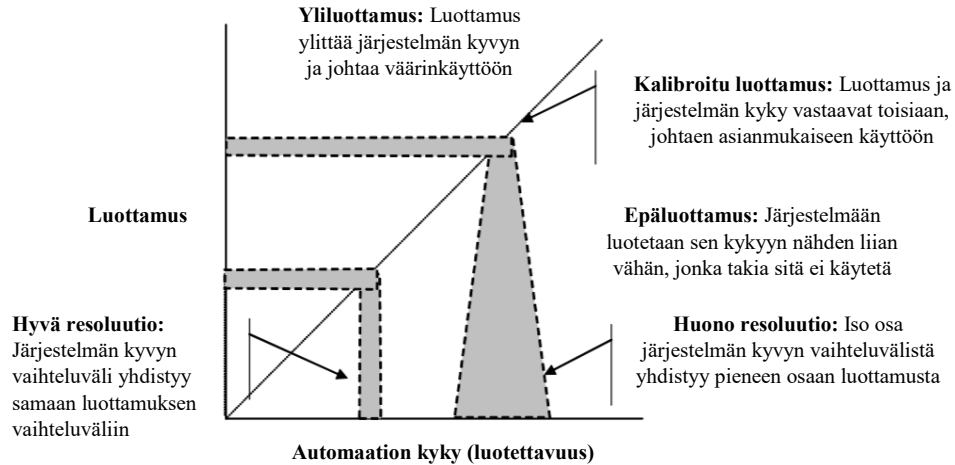
Luottamus on mukana jokapäiväisessä vuorovaikutuksessa ja on keskeinen osa esimerkiksi yhteiskuntaa, kulttuureita, hallituksia ja organisaatioita. Luottamusta ei kuitenkaan ole rajattu vain ihmisten väliseen kanssakäymiseen, vaan se on läsnä myös teknologian saralla. Aivan kuten ihmisten välisissä suhteissa, luottamus määrittelee halukkuutta luottaa myös automatisaatioon esimerkiksi epävarmoissa tilanteissa. Se, kuinka paljon robotteihin luotetaan, vaikuttaa ihmisten ja robottien välisen vuorovaikutuksen laatuun sekä robottien hyväksymiseen ja niiden vakuuttavuuteen (Hancock ym., 2011). Luottamus on esimerkiksi yhteydessä ihmisten halukkuuteen totella robottia sekä hyväksyä sen antamia tietoja (Freed ym., 2007). Robotteja ja tekoälyä kohtaan tunnettu luottamus on siis välttämätöntä niiden täyden hyödyn ja potentiaalain saavuttamista varten. Näin ollen on tärkeää kehittää robotteja

ja tekoälyä, jotka toimivat ja käyttäytyvät tavalla, joka herättää luottamusta ja saa aikaan yhteistyöhalukkuutta.

Koska ihmisten ja robottien tiimejä muodostetaan yhä enemmän, on ihmisten yhteistyöhalukkuus keskeisessä asemassa. Robotit eivät pysty toimimaan itsenäisesti ainakaan vielä ja ihmisten täytyy toimia robottien kanssa. Jotta tiimi voisi saavuttaa tavoitteensa ja suorittaa tehtävänsä, täytyy sen osapuolten (eli tässä tapauksessa ihmisten), pystyä luottamaan toisiin tiimin jäseniin ja heidän toimintaansa sekä kykyynsä suoriutua tehtävästä (Groom & Nass, 2007). Tämä on merkittävää varsinkin tilanteissa, joissa ihmishenkiä on kyseessä. Luottamus tällaisessa tilanteessa on tärkeää, koska se vaikuttaa suoraan ihmisten päätöksentekoon ja käyttäytymiseen eli esimerkiksi päätökseen hyväksyä ja seurata robottien ehdotuksia (Park ym., 2008).

Myös luottamuksen määrä on merkittävä tekijä ihmisten ja robottien tai tekoälyn välisessä vuorovaikutuksessa. Sekä liiallisen että liian vähäisen luottamuksen on huomattu olevan ongelmallista vuorovaikutustilanteissa. Sopivan luottamuksen määrän tukeminen estää automaation väärinkäytön tai sen, ettei sitä käytetä ollenkaan. Kun robottiin luotetaan liikaa, henkilön huomio saattaa siirtyä muualle. Tämä voi johtaa omien tehtävien laiminlyöntiin tai vaikeuksiin päästä tilanteeseen mukaan enää myöhemmin, jos tilanne vaatii puuttumista (Merritt ym., 2012). Leen ja Seen (2004) mukaan tärkeitä käsitteitä luottamuksen määrän tutkimisessa ovat kalibroitu luottamus, yli- ja aliluottamus, hyvä ja huono resoluutio sekä tarkkuus (Kuvio 2). Kuviossa keskellä oleva poikkiviiva kuvaa ihanteellista luottamuksen määrää, jossa luottamuksen määrä ja automaation kyvyt ovat sopusuhteessa.

Kalibroitu luottamus kuvaa henkilön luottamuksen sekä automaation kykyjen välistä suhdetta. Yliluottamuksessa henkilö luottaa automaatioon liikaa eli yli sen kykyjen ja näin ollen on huonosti kalibroitu. Resoluutio taas viittaa siihen, miten tarkasti luottamuksen käsitys erottaa automaation kyvykkyyden tasoja. Kun iso ala automaation kyvyistä sijaitsee pienellä alalla luottamusta, ilmenee huono resoluutio. Tarkkuudella viitataan siihen, kuinka paljon johonkin tiettyyn osa-alueeseen tai taitoon automatisaatiossa luotetaan. Näin ollen hyvä kalibraatio, korkea resoluutio sekä korkea luottamuksen tarkkuus voi vähentää automaation väärinkäyttöä, joka puolestaan auttaa automaation suunnittelussa.



Kuvio 2

Kalibraation, resoluution sekä automaation kyvykkyyden välinen yhteys automaatiota kohtaan tunnetun luottamuksen synnyssä (Lee & See, 2004).

Lee ja Moray (1992) väittävät ihmisten ja automaation välisen luottamuksen muodostuvan automaation suorituksen, menetelmän sekä tarkoituksen perusteella. Mayerin ja kollegoiden (1995) kehittämässä ihmistenvälisessä luottamusmallissa oleva kyky on rinnastettavissa automaation suoritukseen, eli siihen, onko kohde tarpeeksi luotettava, ennustettava sekä kyvykäs. Rehellisyys on taas lähellä menetelmiä, eli sitä, kuinka kohde toimii. Tarkoitus taas viittaa sen tavoitteisiin sekä siihen, suorittaako kohde tehtävänsä tarkoituksen mukaisesti.

Vaikka ihmisten välistä ja ihmisten ja robottien tai tekoälyn välistä luottamusta ei voida suoranaisesti rinnastaa toisiinsa, on näiden luottamustyyppien välillä löydetty kuitenkin yhteneväisyyksiä. Luottamusta tarvitaan molemmissa tapauksissa samanlaisessa asetelmassa: kun jotain vaihdetaan yhteistyössä epävarmuuden vallitessa. Esimerkiksi molemmissa tapauksissa tilannesidonnaiset asenteet, jotka muuttuvat riskin tai epävarmuuden läsnä ollessa, ovat ratkaisevia luottamuksen kannalta (Hoff & Bashir, 2015). Tutkimuksissa on myös löydetty käytettävän samoja neurologisia mekanismeja ihmisten välisessä luottamuspelissä sekä eBay:n nettisivujen luotettavuuden arvioinnissa (Dimoka, 2010). Aivot siis aktivoituvat samankaltaisesti, oli kyseessä luottamus toista ihmistä tai sitten internetsivua kohtaan. Asiaa voi ajatella myös siltä kannalta, että ihmisten luottamus

esimerkiksi automatisaatiota kohtaan edustaa luottamusta asiantuntijoita ja suunnittelijoita kohtaan (Parasuraman & Riley 1997). Näin ollen luottamuksen voi nähdä myös eräänlaisena ihmistenvälisenä luottamuksena, jossa robotti tai tekoäly on mukana välillisesti.

Edellä mainituista tutkimuksista huolimatta ihmisten välisen ja robotteja tai tekoälyä kohtaan tunnetun luottamuksen välillä on mahdollista löytää myös monia eroavaisuuksia. Ensiksi kenties suurimman eron voi huomata luottamuksen kehittämisessä, joka muodostuu ihmistenvälisessä ja ihmisten ja robottien välisessä suhteessa päinvastaisella tavalla. Siinä missä ihmiset lähestyvät toisiaan aluksi varoen ja epävarmasti, robotteja voidaan lähestyä omien ennakkokäsitysten kautta muodostuneiden odotusten takia luottavaisemmin, olettaen niiden olevan täydellisiä (Hoff & Bashir, 2015). Toiseksi ihmistenvälinen luottamus on sosiaalinen prosessi, jossa henkilöt ovat tietoisia toistensa tarkoituksesta ja käyttäytymisestä, joka taas vaikuttaa omaan käyttäytymiseen. Tämä ulottuvuus puuttuu kuitenkin ihmisten ja robottien välisestä vuorovaikutuksesta (Lee & See, 2004). Kolmanneksi iso ero on siinä, että ihmisten ja automaation välisestä luottamuksesta puuttuu tarkoituksellisuus, joka on yksi Leen ja Morayin (1992) luottamuksen synnyn ulottuvuuksista. Ulottuvuuteen liittyy lojaalisuus, hyväntahtoisuus sekä yhdenmukaiset arvot, jotka ovat ihmistenvälisessä luottamuksessa keskeisessä asemassa. Asiaa voi kuitenkin ajatella myös niin, että automaatio välittää sen suunnittelijoiden tarkoituksen, koska se on suunniteltu jotakin tarkoitusta varten.

2.3. Luottamuksen mittaaminen

Luottamuksen mittaaminen on ollut jonkin verran haasteellista sen moniulotteisuuden sekä eri näkökulmien ja lähestymistapojen vuoksi. Luottamusta on esimerkiksi mahdollista tutkia yksilölähteisesti, muun muassa persoonallisuuspiirteiden kautta, kuin myös yhteiskunnan ominaisuutena. Vaikka ihmistenvälistä luottamusta on tutkittu paljon, on luottamuksen mittaaminen ihmisten ja robottien välisessä vuorovaikutuksessa osoittautunut monimutkaisemmaksi ja hankalimmaksi. Suuri ongelma on ollut ihmisten erilaiset käsitykset roboteista ja tekoälystä. Vaikka näillä olisikin omat tekniset määritelmänsä, ihmisten käsitykset ja niihin liitetyt piirteet saattavat poiketa hyvinkin paljon toisistaan (Ray ym., 2008), mikä voi vaikuttaa tutkimustuloksiin merkittävästi.

Luottamuksen mittaamisessa on eniten käytetty erilaisia kyselyjä, joihin on vastattu Likert-asteikolla. Tämän menetelmän käyttöön liittyy kuitenkin jonkin verran ongelmia. Yksi niistä on perinteisesti ollut, ettei kysymyksissä määritellä luottamusta tarkemmin, vaikka ihmisillä saattaa olla hyvinkin erilaiset käsitykset sanan merkityksestä. Usein kysymykset saattavat olla myös liian geneerisiä, jolloin ei välttämättä mainita referenssiryhmää, toimintaa tai panoksia (Ermisch, ym., 2009). Kyselyjen kaltaiset subjektiiviset kertomukset ja raportit eivät aina myöskään välttämättä korreloi itse käyttäytymisen kanssa (Hancock ym., 2011). Lisäksi kyselyt eivät aina mittaa sitä mitä niiden pitäisi. Esimerkiksi Glaeser ja kumppanit (2000) löysivät, että luottamuksen mittaaminen survey-menetelmällä ei mitannut luottamusta vaan enemmänkin luotettavuutta. Tutkimuksen mukaan kyselytulokset sekä luottamuspelin tulokset korreloivat huonosti keskenään. Luottamuspelissä luotettavuus ilmenee rahamäärällä, jonka henkilö saa vastapelaajalta ja luottamus vastapelaajalle annetulla rahamäärällä, mutta kyselyt mittaavat enemmänkin luotettavuutta. Samaan tulokseen päätyivät myös esimerkiksi Lazzarini ja kumppanit (2004), jotka käyttivät samaa tutkimusasetelmaa brasilialaisessa aineistossa.

Muun muassa nämä puutteet kielivät tarpeesta tehdä enemmän objektiivisilla menetelmillä suoritettuja tutkimuksia (Hancock ym., 2011). Kyselyjen ohelle onkin kehitetty joitain kokeellisia menetelmiä, joiden avulla on yritetty välttää survey-menetelmien rajoituksia. Yksi näistä menetelmistä on luottamuspelejä, jota on käytetty laajasti mitattaessa luottamusta. Koska peli mahdollistaa käyttäytymisen tarkkailemisen suoraan, se tarjoaa keinon mitata käyttäytymistä ja yksilöllisiä ominaisuuksia kyselyjä paremmin (Glaeser ym., 2000) ja sisältää yleensä enemmän tietoa panoksista sekä vastapelaajasta (Ermisch ym., 2009). Luottamuspelin avulla pystytään muuntamaan luottamus määrälliseen muotoon ja näin ollen mittaamaan sitä standardisoidusti (Haring ym., 2013). Sitä on kuitenkin kritisoitu muun muassa siitä, ettei vastapelaajalle annettu rahamäärä ilmentäisi luottamusta vaan ihmisten altruismia (Cox, 2004). Tätä väitettä tutkivat muun muassa Brühlhart ja Usunier (2012), jotka päätyivät tutkimuksessaan luottamuspelin mittaavan luottamuksen määrää, eikä altruismi ohjannut ihmisten päätöksiä pelin aikana. Sitä voi siis pitää objektiivisena mittarina, joka arvioi todellista käyttäytymistä (Hancock ym., 2011).

2.4 Luottamukseen vaikuttavat tekijät

Eri tutkijat ovat jakaneet luottamukseen vaikuttavat tekijät eri tavoin. Sanders ja kumppanit (2011) tutkivat luottamukseen vaikuttavia tekijöitä ihmisten ja robottien välisessä vuorovaikutuksessa ja tiimeissä sekä kehittivät ihmisten ja robottien luottamusmallin (engl. A Model of Human-Robot Trust). Lopullinen malli muodostuu kolmesta faktorista (ihmis-, robotti- ja ympäristötekijät), jotka perustuvat empiiriseen tutkimustietoon ja joiden oletetaan vaikuttavan luottamukseen. Malli on kehitetty asiantuntijoiden näkemysten ja laajan kirjallisuuskatsauksen perusteella ja se on saanut tukea myös muissa tutkimuksissa. Robotteihin liittyviä tekijöitä ovat muun muassa robottien ominaisuudet, kuten käyttäytyminen, ennustettavuus, sijainti, persoonallisuus ja sen käyttäytymisen ennustettavuus. Ympäristötekijöitä ovat esimerkiksi kulttuurilliset erot, ajatusmallit, tehtävätyyppi sekä kommunikaatio. Ihmisiin liittyviä tekijöitä taas ovat esimerkiksi taipumus luottaa (Merritt & Ilgen, 2008), itsevarmuus, aikaisempi kokemus ja tilannetietoisuus sekä sosiodemografiset tekijät. Nuorten on esimerkiksi huomattu luottavan robotteihin enemmän kuin vanhempien ihmisten (Scopelliti ym., 2005). Sandersin ja kumppaneiden (2011) mukaan on kuitenkin tärkeää huomioida, että luottamukseen voivat vaikuttaa myös monet ulkopuoliset tekijät, kuten media ja populaarikulttuuri. Media ja populaarikulttuuri saattavat myös selittää robotteja ja tekoälyä kohtaan tunnetun luottamuksen eroja. Siinä missä robotit on elokuvissa usein esitetty hellyttävinä (esim. Star Wars), tekoälyllä on ollut ihmiskunnan kannalta kohtalokkaat seuraukset (esim. Matrix).

Lisäksi Sandersin ja kumppaneiden (2011) tekemän meta-analyysin mukaan ihmisiin perustuvat tekijät ovat mahdollista jakaa kykyihin perustuvaan sekä persoonallisuuteen pohjautuviin tekijöihin. Kykyihin perustuvat tekijät ovat muun muassa sosiodemografiset tekijät, kokemus ja pätevyys, kun taas persoonallisuuteen pohjautuvat tekijät ovat esimerkiksi persoonallisuuspiirteet, robotteja kohtaan olevat asenteet, itsevarmuus sekä taipumus luottamukseen. Analyysissä kuitenkin ihmisiin perustuvat tekijät, kuten persoonallisuus, vaikuttivat vähiten luottamuksen muodostumiseen ihmisten ja robottien vuorovaikutuksessa. Hoff ja Bashir (2015) taas jakavat ihmisten ja automaation väliseen luottamukseen vaikuttavat tekijät kolmeen osa-alueeseen: opittuun, taipumukselliseen sekä olosuhteellisiin tekijöihin. Taipumukselliset tekijät ovat muun muassa kulttuuri, ikä,

sukupuoli sekä persoonallisuuspiirteet. Olosuhteellinen luottamus on riippuvainen kontekstista, jossa vuorovaikutus tapahtuu. Opitussa luottamuksessa taas henkilön aiemmat kokemukset automatisaatiosta ovat oleellisia luottamuksen kehityksessä.

Tutkimusten mukaan luottamus muokkautuu robottien ominaisuuksista (esim. antropomorfismi, turvallisuus, älykkyys) ja ihmisten vuorovaikutukseen vaikuttavista tekijöistä (esim. käytettävyys, sosiaalinen hyväksyttävyys). Varsinkin robottien ulkonäkö on osoittautunut oleelliseksi tekijäksi luottamuksen muodostumisessa. Ihmisten on huomattu luottavan robotteihin, joilla on enemmän ihmismäisiä piirteitä. Robottien ihmismäisyys saattaa esimerkiksi ilmetä silmien tai suun muodolla tai sen käyttäytymisellä (Meyer ym., 2016). Robottien ihmismäisyyden lisääminen saattaa auttaa esimerkiksi tilanteissa, joissa niiden odotetaan tekevän virheitä. Ja päinvastoin, ihmismäisyyden vähentäminen voi olla tarpeen, kun halutaan robotin näyttäytyvän loogisena ja tunteettomana, etäisenä toimijana. Esimerkiksi Lucas ja kumppanit (2014) löysivät tutkimuksessaan, että ihmiset pystyivät ilmaisemaan surun tunnettaan vahvemmin, olivat halukkaampia jakamaan kokemuksiaan, sekä kokivat vähemmän pelkoa, että heidän asiansa tulisi julki, kun he luulivat asioivansa virtuaalisen henkilön kanssa. Ihmisten on huomattu tuntevan myös vähemmän syyllisyyttä asioidessaan koneen kanssa, kun he joutuvat tekemään epäsuotuisia päätöksiä (de Melo & Gratch, 2015). Tämä huomioon ottaen on myös mahdollista, että robotteihin luotettaisiin enemmän kuin abstraktille tasolle jäävään tekoälyyn (Oksanen ym., 2020), koska robottien ulkonäön kautta on mahdollista muodostaa käsitys sen luotettavuudesta. Samoin sosiaalipsykologiassa samankaltaisuuden vetovoimateoria saattaa vaikuttaa siihen, että ihmismäisiin tai samankaltaisen persoonallisuuspiirteen omaaviin robotteihin samaistuttaisiin ja luotettaisiin enemmän (Nass & Lee, 2001). Tämän samankaltaisuuden ja tuttuuden tunteen on huomattu kuitenkin pätevän vain tiettyyn pisteeseen asti, jonka jälkeen robottien ulkonäkö alkaa tuntua epämukavalta tai jopa vahvasti inhottavalta ihmisille. Tätä pistettä on kutsuttu oudoksi laaksoksi (engl. uncanny valley), jonka Mori (1970) esitti ja jota on käytetty myös esimerkiksi populaarikulttuurissa hyödyksi.

Oksanen ja kumppanit (2020) tutkivat robotteja ja tekoälyä kohtaan tunnettua luottamusta luottamuspelin avulla. Tuloksissa ilmeni, että nimimerkillä jdxr894 kutsuttuihin vastapelaajiin luotettiin enemmän kuin Michael nimisiin pelaajiin. Persoonallisuuspiirteistä

avoimuus oli positiivisesti ja tunnollisuus negatiivisesti yhteydessä luottamuksen kanssa. Haringin ja kumppaneiden (2013) tekemässä tutkimuksessa taas ihmismäiselle humanoidille robotille annettiin enemmän rahaa, jos robottia pidettiin älykkäänä ennen pelin alkua. Tutkimuksessa annettiin myös sama määrä rahaa roboteille kuin ihmistä vastaan pelattavissa luottamuspeleissä. Tämän takia tutkijat päättelivätkin, että robotteja pidettiin älykkäinä ja luotettavina vastapelaajina.

Luottamuksen muodostumisesta ihmisten ja robottien tai tekoälyn välisissä suhteissa on siis kehitetty monia teorioita ja siihen liittyvät tekijät on mahdollista jakaa useaan eri luokkaan. Tutkimustulokset vaikuttavat viittaavan kuitenkin siihen, että luottamuksen muodostuminen on monimutkainen prosessi, joka muodostuu monien eri tekijöiden kautta. Tiivistettynä, sekä ihmisten että teknologian ominaisuudet ovat yhteydessä luottamukseen, minkä lisäksi tulee huomioida ympäröivä konteksti. Ihmisiin liittyvät tekijät ovat muun muassa persoonallisuuspiirteet, demografiikka, aiempi kokemus sekä pätevyys, kun taas robottien ominaisuuksista ulkonäkö, nimi, turvallisuus, älykkyys, käyttäytyminen ja sen ennustettavuus sekä persoonallisuus ovat joitain korostettuja faktoreita. Myös ympäristö- tai yhteiskunnalliset tekijät, kuten kulttuurilliset erot, media sekä populaarikulttuuri, ovat nostettu tutkimuksissa esiin.

3 Persoonallisuuspiirteet

3.1 Persoonallisuuden määrittely ja Big Five -piirteet

Persoonallisuus on erä piirteitä, jotka vaikuttavat käyttäytymiseemme ja ajatteluamme (Wood, 2012), toimien motiiveina eri toimintatavoille (Ryckman, 2008). Ne eivät kuvaa pelkästään ihmisten ominaisuuksia, vaan myös henkilöä itseään, näyttäytyen usein eri toimintatilanteissa. Persoonallisuuspiirteet ovat jokseenkin pysyviä ja näin ollen on mahdollista ennustaa ainakin jonkinasteisesti henkilön käyttäytymistä, tunteita ja ajatuksia eri aikoina, eri tilanteissa (Carver & Scheier, 2000; McCrae & Costa, 1994). On olemassa kuitenkin jonkin verran tutkimustuloksia, jotka viittaavat persoonallisuuden myös muuttuvan elämänkaaren aikana. Esimerkiksi Spechtin ja kumppaneiden (2011) tutkimuksessa iän huomattiin vaikuttavan persoonallisuuspiirteiden pysyvyyteen sekä absoluuttisiin arvoihin. Piirteet muuttuivat jopa 30 ikävuoden jälkeen, jota on pidetty persoonallisuuspiirteiden pysyvyyden rajana (Costa & McCrae, 1988). On myös joitain tutkimuksia, joiden mukaan tunnollisuus, sovinnollisuus sekä tunne-elämän tasapaino lisääntyisi iän myötä (Caspi ym., 2003). Hakulisen ja Jokelan (2019) mukaan on kuitenkin tärkeää pitää mielessä, että nämä muutokset vaihtelevat suuresti eri yksilöiden välillä, ja niihin vaikuttavat esimerkiksi yksilön oma valinnat sekä elämän tapahtumat. Persoonallisuus ei muutu siis pelkästään ikääntyessä kypsymisen myötä, vaan ulkoiset tekijät ja muun muassa merkittävät tapahtumat, kuten ensimmäinen työpaikka, lapsen syntymä tai perheenjäsenen kuoleminen, vaikuttavan piirteiden muutoksiin (Specht ym., 2011). Persoonallisuuspiirteet voivat siis muuttua, mutta tämä muutos vaihtelee elämänkulun aikana, ja siihen vaikuttavat moni asia, kuten sosiaaliset odotukset sekä kokemukset. Joka tapauksessa persoonallisuuspiirteiden pysyvyys on ilmennyt eri tutkimuksissa esimerkiksi työelämäikäisillä (Cobb-Clark & Schurer, 2012) ja niiden yleisesti uskotaan ennustavan käyttäytymistä (McAdams & Olson, 2010) sekä kuvaavan pitkäkestoisia eroja ihmisten välillä (Hakulinen & Jokela, 2019).

Persoonallisuuden tutkiminen on hyvin riippuvainen omaksutusta näkökulmasta ja esimerkiksi sosiaalisen oppimisen teoria (Bandura, 1977) sekä psykoanalyttinen ja humanistinen teoria (Lahey, 1992) selittävät ja mittaavat persoonallisuutta. Näiden lisäksi piirreoriat ovat olleet suosittuja selitettäessä persoonallisuutta. Piirre tai ominaisuus on johdonmukainen tai jatkuva tapa ajatella, käyttäytyä ja tuntea, jotka ovat joidenkin

tutkijoiden mukaan persoonallisuuden ytimessä. Piirreteoriat eivät kuitenkaan sinällään ole teorioita eivätkä pyri ymmärtämään niiden syntyä, kuten esimerkiksi psykoanalyttinen teoria, vaan toimivat enemmänkin tapana selittää luonnetta sekä toimintatapoja (Fleeson & Jayawickreme, 2015).

Psykoleksikaalinen lähestymistapa pyrkii selittämään ihmisten välisiä eroja arkipäiväisessä kielessä käytettyjen sanojen kautta. Se on yleispätevä ja validi tapa kuvailla persoonallisuuspiirteitä käyttäen monimuotoista ja laajaa sanastoa (Goldberg, 1981). Allport ja Odbert (1936) käyttivät sanakirjaa löytääkseen persoonallisuutta kuvaavaa sanastoa ja löysivät noin 18000 sanaa, jotka jaettiin neljään ryhmään. Tämän jälkeen Cattell (1956) työsti tätä listaa ja jakoi sekä järjesteli sanat uusiin kategorioihin, luoden 16 persoonallisuusfaktoria (16PF malli). Tupes ja Christal (1992) taas jakoivat faktorianalyyseiden perusteella nämä persoonallisuuspiirteet viiteen eri ulottuvuuteen, joka tunnetaan viiden suuren persoonallisuuspiirteiden mallina eli Big Five -mallina. Big Five -malli on siis teorian sijaan luokittelujärjestelmä, joka jakaa persoonallisuuspiirteet viiteen osa-alueeseen. Tässä mallissa on koottu ihmisten yleisesti käyttämiä sanoja kuvaamaan itseään ja muita, eli mallissa olevat viisi faktoria kuvaavat persoonallisuutta käsitteellisesti hyvin laajalla tasolla. Näin ollen Big Five on läheisesti yhteydessä empiirisiin havaintoihin ja jokainen näistä faktoreista sisältää suuren joukon erillisiä ominaisuuksia (John & Srivastava, 1999). Mallin ulottuvuudet ovat tunnollisuus (engl. conscientiousness), neuroottisuus (neuroticism), ulospäinsuuntautuneisuus (extroversion), sovinnollisuus (agreeableness) sekä avoimuus (openness). Näitä viittä ulottuvuutta voi ajatella jatkumona, jossa jokaisella ulottuvuudella on kaksi ääripäätä ja jossa jokaisen pystyy sijoittamaan johonkin kohtaan jatkumoa. Malli tarjoaa siis keinon tarkastella persoonallisuutta systemaattisesti ja jokainen ulottuvuus sisältää suuren määrän spesifejä tunnusmerkkejä. Taulukossa 1 ovat nämä faktorit esiteltynä lyhyesti.

Taulukko 1

Big Five -mallin ylä- ja alapiirteet (Lönqvist & Tuulio-Henriksson, 2008; McCrae & Costa, 1987).

Yläpiirre	Alapiirre
Neuroottisuus	Ahdistuneisuus Vihamielisyys Masentuneisuus Itsensä tarkkailu Impulsiivisuus Haavoittuvaisuus
Ulospäinsuuntautuneisuus	Sosiaalinen dominanssi Tarmokkuus Itsevarmuus Toiminta Jännityksen etsiminen Positiiviset tunteet
Avoimuus	Mielikuvitus Esteettisyys Avoimuus tunteille Avoimuus toiminnalle Avoimuus ajatuksille Avoimuus arvoille
Sovinnollisuus	Luottamus Rehellisyys Epäitsekkyys Myöntäväisyys Vaativattomuus Herkkämielisyys
Tunnollisuus	Kyvykkyys Järjestelmällisyys Velvollisuudentuntoisuus Saavutuksiin pyrkiminen Itsekuri Harkitsevuus

Sovinnollisuus

Sovinnollisuuteen on liitetty erityisesti piirteet kuten luottamus, altruismi, tottelevaisuus, itsekontrolli ja vaatimattomuus (Patrick, 2011). Tästä ulottuvuudesta matalat pisteet saavat ovat epäluuloisia, yhteistyöhaluttomia, kyynisiä, epäkohteliaita ja ärsyyntyneitä (Feshbach ym., 1996). Koska luottamus on yksi tämän ulottuvuuden tärkeistä piirteistä, sovinnollisten ihmisten on huomattu olevan myös luottavaisempia robotteja sekä tekoälyä kohtaan (Bernotat & Eyssel, 2017).

Tunnollisuus

Tunnolliset henkilöt käyttäytyvät usein normien ja sääntöjen mukaisesti ja piirteet, kuten tarkkaavaisuus, tunnollisuus, energisyys, hellittämättömyys, ahkeruus sekä kunnianhimoisuus ovat liitetty tähän ulottuvuuteen. Matalat pisteet tässä ulottuvuudessa ilmaisevat epäluotettavuutta, laiminlyöntiä, päämäärättömyyttä sekä laiskuutta (Feshbach ym., 1996). Tunnolliset ihmiset välttelevät todennäköisemmin sosiaalista vetelehtimistä (engl. social loafing) tai vapaamatkustamista (engl. free-riding) (Morgeson ym., 2005).

Neuroottisuus

Costa ja McCrae (1992) kuvailee neuroottisuutta käsitteillä, kuten huolestuneisuus, epävarmuus, temperamenttisuus ja itsetietoisuus. Neuroottisuus kuvaillaan siis hyvin usein negatiivisilla käsitteillä ja esimerkiksi Tellegen (1985) lisää neuroottisuuden kuvaukseen muun muassa ahdistuneisuuden, masennuksen, vihan ja häpeän. Neuroottisuus ei sisällä pelkästään negatiivisia piirteitä, vaan myös emotionaalista stressistä johtuvia häiriintyneitä ajatuksia ja käyttäytymistä (Costa & McCrae, 1992). Neuroottiset henkilöt suhtautuvat negatiivisesti moniselitteisiin ja uhkaaviin tilanteisiin (Szalma & Taylor, 2011) ja se on yleisesti yhdistetty huonoon ympäristönmuutoksen sopeutumiseen (Matthews ym., 2003) ja matalaan luottamukseen, koska neuroottisten ihmisten on huomattu ajattelevan potentiaalisia menetyksiä ja riskejä enemmän kuin muiden (Zhang, 2020).

Ulospäinsuuntautuneisuus

Ulospäinsuuntautuneisuus on yhdistetty piirteisiin, kuten positiivisuus, energisyys, puheliaisuus, ystävällisyys ja sosiaalisuus, eli muiden seurassa viihtyminen on erityisesti

tässä ulottuvuudessa keskeisessä asemassa. Matalat pisteet tässä ulottuvuudessa ilmaisevat varautuneisuutta, hiljaisuutta, vakavuutta ja etäisenä pysymistä (Feshbach ym., 1996). Myös dominanssi ja kontrolli ovat tämän ulottuvuuden piirteitä (Driskell ym., 2006). Tämä ulottuvuus on joissain tutkimuksissa yhdistetty positiivisesti robotteja ja tekoälyä kohtaan tunnettuun luottamukseen (Merrit & Ilgen, 2008). Evansin ja Revellen (2008) mukaan erityisesti ulospäinsuuntautuneiden valmius hyväksyä heikkoutta korreloi positiivisesti luottamuksen kanssa.

Avoimuus

Avoimia ihmisiä voi kuvailla luovina ja avaramielisinä sekä uskaliaina. Avoimuutta voi ilmaista tunteissa, teoissa, fantasioissa, ideoissa, estetiikassa ja arvoissa (Costa & McCrae, 1980). Tällä ulottuvuudella matalat arvot saavat ovat muita konservatiivisempia, sovinnaisia sekä emotionaalisesti välinpitämättömiä. Avoimuus on joissain tutkimuksissa liitetty luottamukseen, koska he ovat spontaanimpia hyväksymään uusia ideoita sekä ovat avoimempia uutuuksille ja vaihtelulle (Zhou & Lu, 2011). Tässä ulottuvuudessa matalat pisteet saavat ovat joustamattomia, näkevät epävarmat tilanteet uhkana eivätkä siedä epätietoisuutta.

Big Five -mallia on pitkään käytetty tutkittaessa ja mitattaessa persoonallisuutta ja sen on todettu olevan useissa tutkimuksissa luotettava ja validi mittari (John & Srivastava, 1999), joka on myös osoittautunut universaaliseksi kulttuurienvälisissä tutkimuksissa, jopa hyvin erilaisissa kulttuureissa (McCrae ym., 2005). Persoonallisuuden lähestyminen piirreteorian kautta on tuonut yksimielisyyttä tutkimuskentälle. Aiemmin tutkijoilla oli valittavana ja käytettävänä monia kyselyjä, ilman kattavaa perustelua valinnalle. Ongelmallista oli myös se, että usein näissä kyselyissä käytettiin myös samoja käsitteitä, vaikka ne olivat määritelty eri tavalla (John & Srivastava, 1999). Nämä seikat tekivät tutkimustulosten vertailusta hankalaa ja Big Five -mallin käytön myötä myös tulosten vertailu on helpompaa.

Yksi Big Five -mallin vahvuuksista on, että se mittaa taipumuksellisia persoonallisuuspiirteitä, jotka ovat jokseenkin pysyviä, eivätkä muutu paljon ajan mittaan. Malli tarjoaakin perusteellisen ja teoreettisen lähestymistavan persoonallisuuden tutkimukselle (Meerbeck ym., 2009). Esimerkiksi Santamarian ja Nathan-Robertsin (2017)

tekemässä meta-analyysissä ilmeni Big Five -mallin olevan ylivoimaisesti eniten käytetty malli mitattaessa ihmisten persoonallisuutta ihmisten ja robottien välisessä kanssakäymisessä. Mallin käyttö näin ollen siis mahdollistaa yhtenäisen mittaustavan, joka on vertailtavissa muihin tutkimuksiin.

3.2 Persoonallisuuspiirteiden ja luottamuksen välistä yhteyttä tutkivat aiemmat tutkimukset

Koska persoonallisuus on pysyvä taipumus käyttäytyä, on näin ollen myös luontevaa olettaa, että ihmisten persoonallisuuspiirteet ilmaisevat, osoittavat sekä mahdollisesti ennustavat heidän käyttäytymistään tulevaisuudessa (Robertson & Cooper, 2011). Tutkimusten mukaan persoonallisuuden ja luottamuksen välillä on vahva yhteys ja ihmistenvälistä luottamusta on tutkittu paljon erilaisissa konteksteissa. Müller ja Schwieren (2019) tutkivat erityisesti persoonallisuuspiirteiden asemaa ennustettaessa ihmisten käyttäytymistä luottamuspelissä. Tutkimus toteutettiin laboratoriossa ja tulosten mukaan persoonallisuuspiirteillä on vahva yhteys ensimmäisen pelaajan eli luottavan henkilön käyttäytymiseen. Näistä erityisesti ulospäinsuuntautuneisuuden, sovinnollisuuden ja avoimuuden on huomattu korreloivan positiivisesti ja neuroottisuuden negatiivisesti luottamuksen kanssa ihmisten välisissä suhteissa (Evans & Revelle, 2008; Shikishima ym., 2006).

Vaikka persoonallisuuspiirteiden vaikutusta luottamukseen yleisesti on tutkittu paljon varsinkin ekonomisesta näkökulmasta, tutkimuksia, joissa tarkastellaan niiden vaikutusta robotteja tai tekoälyä kohtaan tunnettuun luottamukseen on vähän. Suurimmassa osassa näistä tutkimuksista on löydetty kuitenkin näyttöä sille, että persoonallisuuspiirteet ovat yhteydessä luottamuksen kanssa ja sillä vaikuttaakin olevan merkittävä rooli ihmisten ja robottien sekä tekoälyn välisessä kanssakäymisessä. Näissä tutkimuksissa Big Five -mallia ja etenkin sisäänpäinsuuntautuneisuus-ulospäinsuuntautuneisuus-ulottuvuutta on käytetty usein tutkittaessa ihmisten ja robottien välistä vuorovaikutusta (Meerbeek ym., 2009; Santamaria & Nathan-Roberts, 2017).

Savela ja kumppanit (2021) tutkivat robottitiimijäsenten käyttämisen vaikutuksia ihmisten ryhmänsisäiseen kuuluvuuden tunteeseen samalla aineistolla, jota on käytetty myös tässä

tutkimuksessa. Kuuluvuuden tunteeseen vaikutti muun muassa persoonallisuus ja ryhmänsisäinen kuuluvuus väheni, kun ryhmän jäsenenä oli robotti. Avoimet, ulospäinsuuntautuneet ja sovinnolliset henkilöt samastuivat vahvasti tiimiin, jossa oli myös robotti, ja neuroottiset ja tunnolliset puolestaan heikommin.

Persoonallisuuspiirteet vaikuttavat yleiseen teknologian hyväksymiseen sekä erityisesti robottien hyväksymiseen (Weiss ym., 2008). Sen on huomattu olevan yhteydessä muun muassa tekoälyyn ja robotteihin suhtautumiseen (Müller & Richert, 2018) sekä proksemiikkaan (välimatka ja tilankäyttö). Ihmisten ja robottien välisessä kanssakäymisessä ihmisten on huomattu myös suosivan robotteja, jotka ovat samankaltaisia itsensä kanssa ja näin ollen myös robotin persoonallisuuden suositaan olevan samankaltainen oman persoonallisuuden kanssa (Craenen ym., 2018). Esimerkiksi Syrdal ja kumppanit (2007) löysivät tutkimuksessaan, että ulospäinsuuntautuneet sekä tunnollisuudessa matalat pisteet saavat osallistujat sallivat robottien lähestyä itseään enemmän kuin sisäänpäin suuntautuneet ja tämä oli huomattavissa varsinkin tilanteessa, jossa ihminen oli kontrollissa (vs. robotti kontrollissa). Yksi ekstroverttien piirteistä on tarve dominoida tai olla kontrollissa, joka mahdollisesti selittää kontrollin tarpeen tutkimuksessa.

Ekstroverttien on huomattu olevan luottavaisempia kuin introverttien automatisoituja kohteita ja robotteja kohtaan, kun taas neuroottiset henkilöt eivät hyväksy automatisoituja kohteita yhtä helposti (Haring ym., 2013). Esimerkiksi Merritt ja Ilgen (2008) tutkivat ihmisten luottamusta ja automatisoitujen kohteiden käyttöä katsomalla, kuinka paljon osallistujat luottivat ja minkälaisia päätöksiä he tekivät koskien matkalaukkujen läpivalaisulaitteen käyttöä. Persoonallisuuspiirteistä vain ulospäinsuuntautuneisuus mitattiin ja tuloksissa havaittiin tilastollisesti merkitsevä yhteys luottamukseen konetta kohtaan. Szalma ja Taylor (2011) taas tutkivat persoonallisuuden ja automatisoituihin kohteisiin suhtautumisen (kuten kohteen luotettavuus) yhteyttä muun muassa käyttäjän tehtävänsuoritukseen ja stressiin ja tulosten mukaan persoonallisuus on yhteydessä automatisoituihin kohteisiin suhtautumiseen. Erityisesti neuroottisuus ennusti huonompaa suoriutumista tehtävästä ja korkeampia stressitasoja.

Salem ja kumppanit (2015) taas tutkivat virheitä tekevän robotin vaikutuksia ihmisten käyttäytymiseen ja yhteistyöhalukkuuteen. Persoonallisuuspiirteistä vain

ulospäinsuuntautuneisuutta ja emotionaalista vakautta, eli matalaa neuroottisuutta mitattiin. Vaikka sekä ulospäinsuuntautuneisuus että emotionaalinen vakaus olivat yhteydessä lämpimään suhtautumiseen robotteja kohtaan, toisin kuin aiemmissa tutkimuksissa, kumpikaan piirre ei vaikuttanut robotteja kohtaan tunnetun luottamuksen kehitykseen objektiivisesti eikä subjektiivisesti arvioitaessa. Toisaalta yksi tutkimuksen rajoituksista oli siihen osallistuneiden ihmisten suhteellisen pieni määrä (N = 40). Myöskään Schaefer ja kumppanit (2012) eivät löytäneet tilastollisesti merkitsevää yhteyttä tutkittaessa persoonallisuuspiirteiden vaikutusta robotteja kohtaan tunnettuun luottamukseen ennen robottien ja ihmisten välistä vuorovaikutusta, näyttämällä robotin kuvia. Tutkijat kuitenkin alleviivaavat tulosten viittaavan luottamuksen määrään vain ennen vuorovaikutusta, eivätkä kerro persoonallisuuden yhteydestä luottamukseen vuorovaikutuksen aikana tai sen jälkeen.

Tekoälyä käytetään hyvin laajasti erilaisissa konteksteissa ja eri tarkoituksia varten. Tekoälyyn liittyen muun muassa digitaalisia assistentteja, sähköistä kaupankäyntiä ja automaattisia autoja on tutkittu mitattaessa persoonallisuuden ja luottamuksen välistä yhteyttä. Esimerkiksi Noahin ja Sethumadhavanin (2019) tutkimuksessa, jossa mitattiin kolmen eri sukupolven tuntemaa luottamusta digitaalisia assistentteja (esim. Siri, Cortana) kohtaan, sovinnollisuus oli ainoa persoonallisuuspiirre, joka oli yhteydessä luottamukseen ja yhteys näkyi van 13–24-vuotiaiden ikäryhmässä. Zhou ja Lu (2011) taas tutkivat ihmisten luottamusta sähköiseen kaupankäyntiin ja tulosten mukaan kaikki persoonallisuusulottuvuudet paitsi tunnollisuus oli yhteydessä luottamukseen mobiilikauppoja kohtaan. Näistä ulospäinsuuntautuneisuudella oli suurin vaikutus, kun taas neuroottiset henkilöt olivat muita epäileväisempiä sekä negatiivisempia. Myös Zhangin ja kumppaneiden (2020) tutkimuksessa, jossa mitattiin luottamusta automatisoituja ajoneuvoja kohtaan, neuroottisuus oli negatiivisesti yhteydessä luottamuksen kanssa ja neuroottiset henkilöt ilmaisivat muita vähemmän halua käyttää automatisoituja ajoneuvoja. Toinen luottamuksen kanssa korreloiva ulottuvuus oli avoimuus ja avoimet henkilöt olivat halukkaampia käyttämään automatisoituja ajoneuvoja.

Persoonallisuuspiirteistä siis ainakin avoimuus, ulospäinsuuntautuneisuus ja sovinnollisuus ovat liitetty positiivisesti ja neuroottisuus negatiivisesti luottamuksen syntyyn ja kehitykseen. Kuitenkin osa tuloksista ovat olleet ristiriitaisia ja lisää tutkimuksia tarvitaan tällä alalla

(Hancock ym., 2011). On tärkeää myös huomata, että tutkimustulokset saattavat vaihdella tutkimusasetelman mukaan ja ihmisten aikaisemmat kokemukset ja taidot vaikuttavat myös merkittävästi. Persoonallisuuspiireet vaikuttavat luottamukseen eri tavalla riippuen tilanteen tai suoritettavan tehtävän ajasta. Esimerkiksi Szalma ja Taylor (2011) huomauttavat, että luottamukseen yhteydessä olevat persoonallisuuspiirteet korreloivat positiivisesti, kun automaation luotettavuudesta ei olla varmoja (esimerkiksi tehtävän suorittamisen varhaisissa vaiheissa) ja kun kohteen luotettavuudesta on enemmän tietoa, persoonallisuuden merkitys vähenee ja tilannekohtaisuus on tärkeämpää. Tulokset ovat myös pitkälti riippuvaisia automaation kohteesta ja suoritettavasta tehtävästä (Szalma & Taylor, 2011) ja näin ollen saattavat selittää ristiriitaiset tulokset alalla.

4 Tutkimuskysymykset ja hypoteesit

Tämän tutkimuksen tavoitteena on selvittää, mitkä persoonallisuuspiirteet vaikuttavat robotteja ja tekoälyä kohtaan tunnettuun luottamukseen, eli luottamuspelissä annettuun rahamäärään. Lisäksi tutkitaan taustamuuttujien yhteyttä luottamukseen sekä luotetaanko robotteihin tai tekoölyyn toista enemmän.

Tutkimuskysymykset ovat siis seuraavat:

- 1) Miten Big Five -persoonallisuuspiirteet ovat yhteydessä robotteja ja tekoälyä kohtaan tunnettuun luottamukseen?
- 2) Eroaako luottamus tekoölyyn luottamuksesta robotteja kohtaan?
- 3) Ovatko sukupuoli, ikä, tulot tai koulutus yhteydessä robotteja ja tekoälyä kohtaan tunnettuun luottamukseen?

Kuten kirjallisuuskatsauksessa todettiin, tutkimusten mukaan persoonallisuuspiirteistä erityisesti ulospäinsuuntautuneisuus (Evans & Revelle, 2008; Haring ym. 2013; Merrit & Ilgen, 2008; Shikishima ym., 2006; Zhou & Lu, 2011) sekä avoimuus (Zhang ym., 2020; Zhou & Lu, 2011) ovat olleet positiivisesti, ja neuroottisuus (Evans & Revelle, 2008; Szalma & Taylor, 2011; Zhang ym., 2020; Zhou & Lu, 2011) negatiivisesti yhteydessä luottamukseen. Näiden tulosten pohjalta on muodostettu seuraavat hypoteesit:

H1: Ulospäinsuuntautuneisuus on positiivisesti yhteydessä robotteihin ja tekoölyyn kohdistuvaan luottamukseen.

H2: Avoimuus on positiivisesti yhteydessä robotteihin ja tekoölyyn kohdistuvaan luottamukseen.

H3: Neuroottisuus on negatiivisesti yhteydessä robotteihin ja tekoölyyn kohdistuvaan luottamukseen.

5 Aineisto ja menetelmät

5.1 Aineisto

Tässä survey-tutkimuksessa käytettiin poikkileikkausaineistoa, joka on kerätty osana monitieteistä *Robotit ja me: vuorovaikutuksen fysiologinen, psykologinen ja sosiaalinen ulottuvuus* -tutkimushanketta (2018–2019). Hanke on rahoitettu Suomen Kulttuurirahaston Pirkanmaan rahaston toimesta ja se pyrkii selvittämään ihmisten suhtautumista robotteihin sekä tutkii ihmisten ja robottien välistä vuorovaikutusta. Konsortioprojektin johtajat ovat Jari Hietanen, Atte Oksanen sekä Veikko Sariola ja hankkeen sosiaalipsykologista osuutta johtaa Atte Oksanen. Sosiaalipsykologian osuuden tutkijoina ovat toimineet Nina Savela ja Rita Latikka.

Aineisto kerättiin huhtikuussa 2019 käyttäen Amazonin Mechanical Turk (MTurk) vastaajapaneelia ja verkkokyselyyn vastasivat Yhdysvalloissa asuvat henkilöt ($N = 1023$). Aineistosta poistettiin 54 vastaajaa, jotka eivät olleet Yhdysvalloista, joten tutkimuksen lopullinen otoskoko oli 969 henkilöä. Näistä henkilöistä 309 kuului tekoäly- ja 325 robottiryhmään.

5.2 Tutkimusetiikka

Osallistuminen perustui vapaaehtoisuuteen ja ennen aineistonkeruuta osallistujille annettiin tietoa tutkimuksesta sekä kerrottiin, että siitä voisi jättäytyä pois missä vaiheessa tahansa. Osallistujat antoivat suostumuksensa osallistua tutkimukseen sekä aineiston jatkokäsittelyyn. Tutkimus sai Tampereen alueen ihmistutkimuksen eettiseltä neuvottelukunnalta (2018) myönteisen lausunnon, jonka mukaan tutkimushanke ei sisältänyt eettisiä ongelmia. Itse allekirjoitin aineistoon liittyvän käyttöluvasopimuksen ennen aineiston vastaanottamista, jossa sitouduin käyttämään aineistoa vain Pro gradun kirjoittamisen ajan. Samalla sitouduin huolehtimaan aineiston tietoturvallisuudesta sekä etten luovuttaisi tietoja ulkopuolisille.

5.3 Muuttujat

Riippuvana eli selitettävänä muuttujana tutkimuksessa oli roboteille ja tekoälylle annettu rahamäärä, jolla mitattiin luottamusta ja riippumattomana eli selittävänä muuttujana Big Five -mallin viisi persoonallisuusulottuvuutta avoimuus, tunnollisuus, neuroottisuus, ulospäinsuuntautuneisuus ja sovinnollisuus sekä taustamuuttujat.

5.3.1 Luottamus

Luottamusta robotteihin ja tekoölyyn mitattiin käyttämällä luottamuspelejä, jossa osallistujia pyydettiin kuvittelemaan vuorovaikutustilanne tuntemattoman henkilön, robotin tai tekoölyn kanssa. Osallistujia pyydettiin kuvittelemaan tilanne, jossa vastassa oli henkilö, robotti tai tekoöly ja jolla oli joko yleinen ihmisnimi (Michael) tai nimimerkki (jdrx894). Heille siis määrättiin satunnaisesti yksi vastapelaaja näistä kuudesta ryhmästä. Osallistujille annettiin 1000 \$ kuvitteellista rahaa ja kerrottiin, että hän saa joko pitää rahan kokonaan itsellään tai antaa haluamansa summan vastapelaajalle. Osallistujille kerrottiin, että annettu summa kolminkertaistettaisiin ennen vastapelaajalle antamista ja vastapelaaja saisi halutessaan antaa tästä rahasummasta osallistujalle haluamansa määrän takaisin. Osallistujat olivat tietoisia luottamuspeleiden hypoteettisuudesta. Luottamuspelejä on yleisesti käytetty menetelmä, joka on alun perin lähtöisin talouden ja sijoittamisen alalta (Camerer, 2003; Johnson & Mislin, 2011).

Kuvailevien tunnuslukujen perusteella ryhmät olivat jakautuneet tasaisesti taustamuuttujien ja persoonallisuuspiirteiden suhteen. Koska analyysissä robottiryhmien ja tekoölyryhmien sisällä ei ollut eroja nimen perusteella, nämä ryhmät yhdistettiin yhdeksi robotti- ja yhdeksi tekoölyryhmäksi eli analyysissä oli kaksi selitettävää muuttujaa. Robotti- ja tekoölyryhmälle annetun rahamäärän kuvailevat tunnusluvut ovat esitettynä taulukossa 2. Tulosten perusteella robottiryhmälle annettiin hieman enemmän rahaa ($ka = 500,19$, $kh = 17,54$) kuin tekoölyryhmälle ($ka = 470,80$, $kh = 17,72$).

5.3.2 Persoonallisuus

Persoonallisuusmittarina käytettiin lyhennettyä Big Five Inventory BFI-S kyselylomaketta (Lang ym., 2011), joka sisältää 15 seitsenasteisella Likert -asteikolla mitattavaa kysymystä.

Väittämiin vastataan valitsemalla arvot 1 (vahvasti eri mieltä) ja 7 (vahvasti samaa mieltä) välillä. Jokainen ulottuvuus mitataan kolmen kysymyksen avulla ja ulottuvuuksista saatavat arvot vaihtelevat 3–21 välillä. Mitä korkeammat pisteet saadaan, sitä vahvempi henkilön kyseinen persoonallisuuspiirre on. Lyhennetyn persoonallisuusmittarin käyttö mahdollistaa tehokkaan ja luotettavan tavan mitata persoonallisuutta ja onkin laajasti käytetty (Lang ym., 2011). Tutkimusten mukaan lyhennetyn lomakkeen käyttö saattaa kuitenkin myös korottaa virheiden määrää ja alentaa tulosten luotettavuutta. Tämän on huomattu kuitenkin koskevan usein mittareita, joissa persoonallisuutta on mitattu vain yhden tai kahden kysymyksen avulla (Credé ym., 2012) ja jokaisen ulottuvuuden mittaamista vähintään kolmen kysymyksen avulla on pidetty luotettavana menetelmänä (Gagné & Hancock, 2006).

Persoonallisuuspiirteiden keskiarvot, keskihajonnat, vaihteluväli sekä Cronbachin alfat ovat esitettynä taulukossa 2. Tulosten perusteella korkein keskiarvo molemmissa ryhmissä on tunnollisuus -ulottuvuudella (robottiryhmässä $ka = 16,36$, $kh = 0,19$ ja tekoälyryhmässä $ka = 16,00$, $kh = 0,19$) ja matalin neuroottisuudella (robottiryhmässä $ka = 10,58$, $kh = 0,29$ ja tekoälyryhmässä $ka = 11,13$, $kh = 0,28$), mutta tilastollista merkitsevyyttä ei ole mahdollista arvioida näiden tulosten perusteella.

Mittarin sisäistä luotettavuutta eli reliabiliteettia tarkasteltiin Cronbachin alfan (α) avulla, joka on käytetyin menetelmä mittareiden luotettavuuden arvioinnissa. Sisäinen luotettavuus kertoo, mittaavatko mittarin kysymykset samaa asiaa ja mitä suurempi arvo on, sitä yhtenäisempi myös mittari on (Tavakol & Dennick, 2011). Alfa-arvoja, jotka ovat alle 0,6, pidetään yleensä huonoina ja näillä mittareilla on matala luotettavuus. Tässä tutkimuksessa ulottuvuuksien Cronbachin alfat vaihtelivat hyvän ja hyväksyttävän välillä ($\alpha = 0,62$ – $0,87$) molemmissa ryhmissä (taulukko 3) (Taber, 2018).

5.3.3 Taustamuuttujat

Taustamuuttujina tutkimuksessa ovat kategorisina muuttujina sukupuoli, kotitalouden vuotuiset bruttotulot, työllisyystilanne ja koulutus sekä jatkuvana muuttujana ikä. Sukupuoli koodattiin dummy-muuttujaksi, jossa nainen sai arvon 0 ja mies arvon 1. Naisia koko aineistossa oli 51,15 % ($n = 488$) ja miehiä 48,85 % ($n = 466$).

Taulukko 2

Kategoristen muuttujien kuvailevat tunnusluvut robotti- (n = 325) ja tekoälyryhmissä (n = 309).

	<i>n</i>		<i>%</i>	
Muuttujat	Robotti	Tekoäly	Robotti	Tekoäly
Sukupuoli				
Nainen	155	162	47,69	52,43
Mies	170	147	52,31	47,57
Työllisyystilanne				
Opiskelija	9	10	2,77	3,24
Työskentelee koko/osa- aikatyössä	278	267	85,54	86,41
Työtön	38	32	11,69	10,36
Kotitalouden vuotuiset tulot				
Alle 35,000 \$	87	76	26,77	24,60
35,000 \$-154,999 \$	223	217	68,62	70,23
155,000 \$ tai yli	15	16	4,62	5,18
Koulutus				
Lukio tai alempi	45	25	13,85	8,09
Korkeakoulu	234	227	72,00	73,46
Maisteri	46	57	14,15	18,45

Kotitalouden vuotuiset bruttotulot koodattiin kolmiluokkaiseksi, jonka kategoriat ovat ”alle 34 999 \$”, ”35 000 \$–154 999 \$” ja ”yli 154 999 \$”. Osallistujien vuositulot vaihtelivat suurimmalla osalla 35 000 \$ ja 154 999 \$ välillä. Koulutusta tarkasteltiin myös kolmessa luokassa: ”toinen aste tai alempi tutkinto” (peruskoulu tai lukio/ammattillinen koulu), ”korkeakoulutus” (alempi korkeakoulututkinto) sekä ”maisteri tai korkeampi tutkinto” (ylempi korkeakoulututkinto tai tohtorintutkinto) ja suurin osa osallistujista oli käynyt korkeakoulun (73,79 %). Työllisyystilanne koodattiin uudelleen kolmena kategoriana: ”opiskelijat”, ”koko- tai osa-aikatyölliset” sekä ”työttömät”. Suurin osa kyselyyn vastanneista olivat osallistumishetkellä koko- tai osa-aikatöissä (87,19 %). Kategoristen muuttujien tunnusluvut ovat kuvattuna taulukossa 2 ja jatkuvien taulukossa 3.

Taulukko 3

Jatkuvien muuttujien kuvailevat tunnusluvut

Muuttujat	<i>n</i>	<i>ka</i>	<i>kh</i>	Vaihteluväli	α
Ikä robottiryhmä	325	37,73	0,64	15–71	
Ikä tekoälyryhmä	309	36,58	0,60	16–71	
Robottiryhmän antama rahamäärä	325	500,19	17,54	0–1000	
Tekoälyryhmän antama rahamäärä	309	470,80	17,72	0–1000	
Robottiryhmän persoonallisuuspiirteet					
Neuroottisuus	325	10,58	0,29	3–21	0,87
Ulospäinsuuntautuneisuus	325	11,37	0,26	3–21	0,85
Avoimuus	325	15,37	0,21	3–21	0,79
Sovinnollisuus	325	15,55	0,20	3–21	0,64
Tunnollisuus	325	16,36	0,19	3–21	0,73
Tekoälyryhmän persoonallisuuspiirteet					
Neuroottisuus	309	11,13	0,28	3–21	0,81
Ulospäinsuuntautuneisuus	309	11,30	0,27	3–21	0,81
Avoimuus	309	15,28	0,22	3–21	0,79
Sovinnollisuus	309	15,07	0,21	3–21	0,62
Tunnollisuus	309	16,00	0,19	3–21	0,67

5.4 Analyysimenetelmät

Analyysimenetelminä käytettiin kuvailevia tilastollisia analyyseja, kuten frekvenssejä, keskiarvoja, keskihajontaa ja prosenttiosuuksia, sekä lineaarista regressioanalyysia. Aineisto analysoitiin käyttäen StataIC-16-ohjelmaa.

5.4.1 Kuvailevat analyysit

Riippumattomien otosten t-testi

Riippumattomien otosten t-testillä on mahdollista tarkastella ryhmien keskiarvojen välisiä eroja. Nollahypoteesina silloin on, että molempien muuttujien keskiarvot ovat yhtä suuria ryhmissä. Vastahypoteesina taas on, että keskiarvot ovat erisuuret (Nummenmaa, 2008).

Testillä tarkasteltiin, annettiinko robotti- tai tekoälyryhmässä tilastollisesti merkitsevästi enemmän rahaa.

Korrelaatio

Korrelaatio on kahden muuttujan välinen tilastollinen riippuvuus. Korrelaatiokerroin kertoo kahden tilastomuuttujan välisen lineaarisen riippuvuuden voimakkuudesta. Jos molemmat muuttujat kasvavat samanaikaisesti, on kyseessä positiivinen korrelaatio, ja jos molemmat vähenevät, korrelaatio on negatiivinen. Yleisimmin korrelaatiota kuvaamaan käytetään Pearsonin tulomomenttikorrelaatiokerrointa (r), joka on mahdollista laskea vähintään välimatka-asteikollisille muuttujille. Korrelaatiokertoimet vaihtelevat -1 ja 1 välillä, jossa ± 1 viittaa täysin lineaariseen yhteyteen ja alle $\pm 0,3$ kerrointa pidetään heikkona lineaarisena yhteytenä (Nummenmaa, 2008).

5.4.2 Lineaarinen regressioanalyysi

Selittävien muuttujien yhteyttä selitettävään muuttujaan tarkasteltiin käyttäen lineaarista regressioanalyysiä. Regressioanalyysi on vahva metodi tarkasteltaessa muuttujien välistä yhteyttä sekä ennustettaessa tutkittavien mittaustuloksia jollain muuttujalla. Tätä varten luodaan matemaattinen esitys, joka kuvaa muuttujien välisiä yhteyksiä (Gkioulekas & Papageorgiou, 2019). Lineaarisen regressioanalyysin etuna on mahdollisuus tarkastella usean selittävän muuttujan vaikutusta selitettävään muuttujaan samanaikaisesti, sekä vakioita useita sekoittavia muuttujia tutkittaessa selittävän ja selitettävän muuttujan välistä yhteyttä. On kuitenkin tärkeää huomioida, että regressioanalyysi ei kuitenkaan kerro muuttujien välisestä syy-seuraus-suhteesta, vaikka muuttujien välillä olisikin vahva yhteys (Nummenmaa, 2008).

Lineaarinen regressioanalyysi tarkastelee muuttujien välistä lineaarista yhteyttä. Yksi regressioanalyysin tehtävistä on arvioida regressiomallin tuntemattomat parametrit eli mallin sopeuttaminen aineistoon. Regressiomalli muodostetaan jäännöstermien avulla, joka on havaitun selitettävän muuttujan arvo vähennettynä selitettävän muuttujan ennustetusta arvosta. Jäännöstermit ovat siis selitettävän muuttujan vaihtelussa se osa, jota regressiomalli

ei pysty selittämään. Näin ollen mitä suurempi jäännöstermi on, sitä huonommin muuttujan arvo on ennustettu mallin avulla (Nummenmaa, 2008).

Standardoimattomat (B) sekä standardoidut regressiokertoimet Beta (β), jotka voivat olla sekä positiivisia että negatiivisia, kuvastavat, kuinka paljon jokainen selittävä muuttuja selittää selitettävän muuttujan vaihtelusta ja miten ne vaikuttavat mallin toimivuuteen. Regressiokertoimen vakioiminen eli standardoiminen mahdollistaa muuttujien suuruuksien keskinäisen vertailun ja näin ollen suhteellisen selityskyvyn vertailemisen. Regressiokerroin standardoidaan niin, että keskiarvo on 0 ja keskihajonta 1 (Nummenmaa, 2008).

Kun regressiomalli on luotu, tarkastellaan mallin sopivuutta, jolloin on mahdollista muokata mallia lisäämällä tai poistamalla muuttujia. Linearisessa regressioanalyysissä on siis mahdollista käyttää useampaa selittäjää, joka toimii samalla tavalla kuin yhden selittäjän malli. Useamman selittäjän käyttö parantaa usein mallinnuksen tarkkuutta eli lisää mallin selitysstettä. Hyvässä regressiomallissa jäännöseliösumma on pieni ja selitysstaste korkea. Korrelaatiokertoimen toinen potenssi eli R^2 on regression selitysstaste, joka kertoo prosenttimäärän siitä, kuinka paljon muuttujat selittävät toistensa arvojen vaihtelusta (Nummenmaa, 2008).

Regressiomalli on mahdollista muodostaa useista eri vaihtoehdoista. Askeltavassa regressioanalyysissä, jota kutustaan myös tilastolliseksi regressioksi (stepwise), on mahdollista lisätä ja poistaa muuttujia mallista vaiheittain. Näiden järjestys päätetään tutkimalla F-testin arvoja ja korkein F-testin arvo verrataan valittuun tai oletusarvoiseen F-testin arvoon. Muuttujia lisätään ja poistetaan yksi kerrallaan, kunnes saadaan tulokseksi tilastollisesti kaikkein parhaiten sopiva malli aineistoon. Tässä tutkimuksessa regressiomalli muodostettiin tutkijan teoreettisesti määrittelemällä mallilla eteenpäin askeltavana mallina. Etenevästi muuttujia lisättäessä sijoitetaan malliin vain vakio, jonka jälkeen lisätään järjestyksessä muut muuttujat, jotka kasvattavat mallin selitysstettä (Chatterjee & Simonoff, 2013).

Tässä tutkimuksessa tehtiin kolme eri mallia. Ensimmäisessä mallissa tutkittiin kunkin persoonallisuuspiirteen yhteyttä selitettävään muuttujaan itsessään, toisessa mallissa tutkittiin kaikkien persoonallisuuspiirteiden yhteisvaikutusta selitettävään muuttujaan ja kolmannessa mallissa olivat kaikki muuttujat mukana.

Lineaarisen regressioanalyysin ehdot

Lineaarista regressioanalyysiä voi käyttää, kun selitettävä muuttuja on numeerinen eli vähintään välimatka-asteikollinen. Regressioanalyysin käyttöön liittyy tiettyjä oletuksia ja ehtoja, jotka pitää täyttää. Kenties tärkein näistä ehdoista on, että muuttujien välillä on lineaariset yhteydet. Tätä on helpointa tutkia muuttujien sirontamatriisin avulla, jota käytettiin tässä tutkimuksessa. Toinen analyysimenetelmän oletuksista on, että selittävät muuttujat eivät ole yhteydessä toisiinsa eli ovat kollinearittomia. Tämä oletus varmistaa, ettei malliin tule selittäjäksi sama vaihtelu useampaan kertaan, jonka takia periaatteessa malliin hyvin sopiva muuttuja saattaisi vaikuttaa huonolta selittäjältä (Chatterjee & Simonoff, 2013). Tässä tutkimuksessa multikollineaarisuus mitattiin tarkastelemalla VIF-arvoja (engl. variance inflation factor), eikä multikollineaarisuutta ilmennyt (liitteet A ja B). Vaikka arvoille ei olekaan asetettu tarkkaa rajaa, yli $VIF > 10$ viittaa jo melko korkeaan multikollineaarisuuteen (Hair ym., 1998).

Regressioanalyysissä myös oletetaan, että mallin jäännöstermit ovat homoskedastisia, eli että jäännöstermien varianssi on vakio. Tämäkin on mahdollista tarkastella esimerkiksi jokaisen selittävän muuttujan kohdalla sirontamatriisin avulla. Tässä tutkimuksessa jäännöstermien eli residuaalien heteroskedastisuus tarkistettiin Breusch-Pagan / Cook-Weisberg testillä. Testin tulos ei ollut merkitsevä, joten residuaalit olivat homoskedastisia. Aineisto ei myöskään saa olla aikariippuvainen, koska käytetyt muuttujat ovat usein riippuvaisia toisistaan vuosien välillä, joka vaikuttaa analyysien luotettavuuteen. Toisin sanoen havaintojen jäännöstermit eivät saa korreloida keskenään (Chatterjee & Simonoff, 2013). Myös otoskolla ja normaalijakaumalla on jonkin verran merkitystä suorittaessa lineaarista regressioanalyysiä. Usein suhteellisen pieni otoskoko, esimerkiksi muutaman selittävän muuttujan analyysissä 50–100 havaintoa, voi olla riittävä määrä. Muiden oletusten ollessa voimassa regressioanalyysi toimii siis hyvin myös pienellä otoskolla. Myös normaalijakaumaoletus on voimassa regressioanalyysissä, mutta suurilla otoskoilla tämä ei ole niin merkittävä. Lopuksi poikkeavat havainnot (outlier) saattavat vaikuttaa regressioanalyysin tuloksiin, joten aineiston tarkastelu myös niiden osalta ennen analyysiä on suositeltua (Nummenmaa, 2008).

6 Tulokset

6.1 Kuvailevat tulokset

Tuloksissa ilmeni, että osallistujat antoivat enemmän rahaa roboteille ($ka = 500,19$, $kh = 17,54$), kuin tekoälylle ($ka = 470,80$, $kh = 17,72$). Riippumattomien otosten t-testin avulla tarkasteltiin, oliko tämä ero tilastollisesti merkitsevä. T-testin perusteella annetut rahamäärät eivät eronneet toisistaan tilastollisesti merkitsevästi, $t(632) = 1,18$, $p > 0,05$. Muuttujien väliset korrelaatiot robottiryhmässä ovat esitettyinä taulukossa 4 ja tekoälyryhmässä taulukossa 5. Kummassakaan taulukossa ei ilmene kovin korkeita korrelaatioita. Taulukossa 4 korkeimmat korrelaatiot ilmenevät esimerkiksi sovinnollisuuden ja tunnollisuuden sekä neuroottisuuden ja tunnollisuuden välillä. Taulukosta 4 myös ilmenee, että koulutuksen ja avoimuuden välillä on tilastollisesti merkitsevä, heikko lineaarinen yhteys ($r = 0,15$, $p < 0,001$). Ikä korreloi tilastollisesti merkitsevästi ja positiivisesti, vaikkakin heikosti, tunnollisuuden ($r = 0,19$, $p < 0,001$) ja sovinnollisuuden ($r = 0,17$, $p < 0,001$), ja negatiivisesti neuroottisuuden ($r = -0,19$, $p < 0,001$) kanssa. Taulukosta 5 ilmenee, että sukupuoli on negatiivisesti ja heikosti, mutta tilastollisesti merkitsevästi yhteydessä tunnollisuuden ($r = -0,12$, $p < 0,01$) ja sovinnollisuuden ($r = -0,24$, $p < 0,001$) kanssa. Vanhempi ikä oli tilastollisesti merkitsevästi yhteydessä tunnollisuuden ($r = 0,35$, $p < 0,001$), sovinnollisuuden ($r = 0,29$, $p < 0,001$), neuroottisuuden ($r = -0,17$, $p < 0,001$) sekä sukupuolen ($r = -0,15$, $p < 0,01$) kanssa.

Taulukko 4

Muuttujien väliset korrelaatiot robottiryhmässä (n = 325)

Muuttujat	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Luottamus	1.000										
2. Avoimuus	0.10*	1.000									
3. Tunnollisuus	-0.01	0.28***	1.000								
4. Ulospäin-suuntautuneisuus	0.02	0.21***	0.09*	1.000							
5. Sovinnollisuus	-0.02	0.29***	0.35***	0.19***	1.000						
6. Neuroottisuus	-0.05	-0.03	-0.39***	-0.33***	-0.29***	1.000					
7. Sukupuoli	0.09*	-0.08	-0.07	0.08	-0.07	-0.16	1.000				
8. Koulutus	0.60	0.15***	0,04	0.06	-0.02	0.01	0.03	1.000			
9. Ikä	0.03	0.06	0.19***	0.08	0.17***	-0.19***	-0.06	0.07	1.000		
10. Tulot	0.02	-0.07	0.13**	0.16***	0.07	-0.12***	0.06	0.17***	-0.07	1.000	
11. Työ	-0.02	-0.02	-0.14***	-0.11**	-0.00	0.05*	-0.07	-0.10*	0.12**	-0.14**	1.000

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Taulukko 5

Muuttujien väliset korrelaatiot tekoälyryhmässä (n = 309)

Muuttujat	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. Luottamus	1,000										
2. Avoimuus	0,15***	1.000									
3. Tunnollisuus	-0,09	0.24***	1.000								
4. Ulospäin-suuntautuneisuus	0,01	0.26***	0.10*	1.000							
5. Sovinnollisuus	-0,01	0.23***	0.43***	0.11*	1.000						
6. Neuroottisuus	-0.09	-0.08	-0.39***	-0.20***	-0.25***	1.000					
7. Sukupuoli	0.03	-0.05	-0.12**	-0,01	-0.24***	-0.10*	1.000				
8. Koulutus	0.10*	-0.05	-0,00	0.12**	-0.09*	-0,06	0,11**	1.000			
9. Ikä	0.026	0.08	0.35***	0.08	0.29***	-0.17***	-0,15**	0,05	1.000		
10. Tulot	0.00	0,03	0,08	0,05	0.05	-0.09	0,03	0.17***	0.06	1.000	
11. Työ	-0.03	0,01	0,04	-0.06	0,16***	0.02	-0.10*	-0.08	0.19***	-0.01	1.000

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

6.2 Lineaarisen regressioanalyysin tulokset

Lineaarinen regressioanalyysi suoritettiin erikseen robotti- sekä tekoälyryhmille kolmessa eri mallissa. Muuttujia lisättiin malleihin vaiheittain ja samalla tarkasteltiin mallien muutoksia. Ensimmäisessä mallissa tarkasteltiin jokaisen persoonallisuuspiirteen yhteyttä luottamukseen itsessään, toisessa mallissa kaikkien persoonallisuuspiirteiden vaikutusta luottamukseen yhdessä sekä kolmannessa persoonallisuuspiirteiden ja kontrollien vaikutuksia luottamukseen. Tässä osiossa raportoidaan taulukkomuodossa ainoastaan molemmissa ryhmissä korkeimman selitysasteen omaava Malli 3, kun taas muut mallit raportoidaan kirjallisesti. Lineaarisen regressioanalyysin tulokset ovat esitettynä taulukossa 6.

Ensimmäisessä mallissa mikään persoonallisuuspiirteistä ei ennustanut roboteille annettua rahamäärää tilastollisesti merkitsevästi, kun taas tekoälyryhmässä ainoastaan avoimuus oli yhteydessä annetun rahamäärän kanssa ($\beta = 0,15, p < 0,01$), mallin selitysasteen ollessa 2 prosenttia. Toisessa mallissa avoimuus oli sekä robottiryhmässä ($\beta = 0,13, p < 0,05, R^2 = 0,00$) että tekoälyryhmässä ($\beta = 0,18, p < 0,01, R^2 = 0,03$) tilastollisesti merkitsevästi yhteydessä koettuun luottamukseen.

Kolmannen mallin robottiryhmässä persoonallisuuspiirteistä ainoastaan avoimuus oli tilastollisesti merkitsevästi yhteydessä annetun rahamäärän kanssa ($\beta = 0,14, p < 0,05$). Myöskään mikään taustamuuttujista ei ollut tilastollisesti yhteydessä roboteille annetun rahamäärän kanssa. Malli selittikin selitettävän ilmiön vaihtelusta 0 prosenttia, eikä ollut tilastollisesti merkitsevä.

Tekoälyryhmässä persoonallisuuspiirteistä avoimuus ($\beta = 0,18, p < 0,01$) oli positiivisesti ja tunnollisuus negatiivisesti ($\beta = -0,17, p < 0,05$) yhteydessä selitettävään muuttujaan. Taustamuuttujista ikä ($\beta = 0,13, p < 0,05$) oli tilastollisesti yhteydessä tekoälylle annetun rahamäärän kanssa, mikä tarkoittaa, että iän kasvaessa myös tekoälylle annettu rahamäärä kasvoi. Malli oli tilastollisesti merkitsevä ($p < 0,05$) ja mallin selitysaste oli 4 prosenttia. Tulosten perusteella sukupuoli, koulutus, kotitalouden vuotuiset tulot tai työllisyystilanne ei siis vaikuttanut roboteille tai tekoälylle annettuun rahamäärään.

Taulukko 6

Lineaarisen regressioanalyysin tulokset koskien persoonallisuuspiirteitä ja taustamuuttujia sekä robotille/tekoälylle annettua rahamäärää

Muuttuja	Robotit			Tekoäly		
	B	SE B	β	B	SE B	β
Persoonallisuuspiirteet						
Avoimuus	36,75*	16,23	0,14	43,62**	14,81	0,18
Tunnollisuus	-20,10	19,05	-0,07	-46,27*	19,41	-0,17
Ulospäin-						
suuntautuneisuus	-7,73	12,35	-0,53	-7,28	11,96	-0,04
Sovinnollisuus	-12,26	17,08	-0,05	15,55	17,04	0,06
Neuroottisuus	-13,11	12,34	-0,07	16,25	12,13	0,09
Sukupuoli						
Nainen	0,00	0,00	0,00	0,00	0,00	0,00
Mies	54,95	36,43	0,09	15,99	37,10	0,03
Koulutus						
Lukio tai alempi	0,00	0,00	0,00	0,00	0,00	0,00
Korkeakoulu	-,42	53,08	-0,00	71,75	65,45	0,10
Maisteri tai						
korkeampi	39,40	69,24	0,04	110,34	75,61	0,14
Ikä	1,16	1,64	0,04	3,94*	1,82	0,13
Vuotuiset tulot						
Alle 35000 \$	0,00	0,00	0,00	0,00	0,00	0,00
35000–149999 \$	53,36	42,35	0,08	47,59	42,20	0,07
Yli 150000 \$	-55,73	91,43	-0,04	-75,10	85,57	-0,05
Työllisyystilanne						
Opiskelija	0,00	0,00	0,00	0,00	0,00	0,00
Koko/osa-						
aikatyö	-48,84	109,90	-0,05	2,04	101,55	0,00
Työtön	-62,68	121,09	-0,06	-56,93	114,00	-0,06
Havaintojen määrä		325		309		
Mallin		0,48		0,02		
merkitsevyys (p)						
R^2		0,04		0,08		
Adj. R^2		-0,00		0,04		

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

7 Pohdinta

7.1 Pohdinta

Tämän tutkimuksen tavoitteena oli tarkastella, vaikuttavatko persoonallisuuspiirteet robotteja tai tekoälyä kohtaan tunnettuun luottamukseen. Persoonallisuuspiirteiden yhteyttä luottamukseen mitattiin käyttäen hierarkkista lineaarista regressioanalyysiä, jonka avulla tehtiin kolme eri mallia. Samalla tutkittiin taustamuuttujien, kuten iän, sukupuolen, työn ja koulutuksen yhteyttä robotteja ja tekoälyä kohtaan tunnettuun luottamukseen sekä sitä, miten luottamus robotteihin eroaa luottamuksesta tekoälyyn.

Tulosten perusteella mikään persoonallisuuspiirteistä ei selitä vahvasti robotteja tai tekoälyä kohtaan tunnettua luottamusta. Robottiryhmän ensimmäisessä mallissa mikään muuttujista ei ollut tilastollisesti merkitsevästi yhteydessä koetun luottamuksen kanssa, kun taas toisessa ja kolmannessa mallissa avoimuuden ja luottamuksen välillä löydettiin yhteys, mutta mallit eivät olleet tilastollisesti merkitseviä. Tekoälyryhmässä avoimuus oli yhteydessä luottamukseen kaikissa malleissa ja näiden lisäksi kolmannessa mallissa löydettiin myös negatiivinen yhteys tunnollisuuden ja positiivinen yhteys iän kanssa, eli iän kasvaessa myös tekoälylle annettu rahamäärä kasvoi. Vaikka tekoälyryhmän malli oli tilastollisesti merkitsevä, sen selitysaste oli matala. Tulosten perusteella sukupuolella, koulutuksella, kotitalouden vuotuisilla tuloilla tai työllisyystilanteella ei ollut yhteyttä roboteille tai tekoälylle annettuun rahamäärään.

Tutkimuksessa asetetut hypoteesit 1 ja 3, eli ulospäinsuuntautuneisuuden sekä neuroottisuuden positiivinen yhteys luottamuksen kanssa ja sen ennustaminen, kumottiin, eivätkä nämä persoonallisuuspiirteet olleet selittäviä tekijöitä missään malleista. Aiemmissa tutkimuksissa ulospäinsuuntautuneisuuden ja matalan neuroottisuuden on huomattu olevan yhteydessä luottamuksen kanssa ainakin ihmistenvälisessä vuorovaikutuksessa (Evans & Revelle, 2008; Shikishima ym., 2006). Myös Zhou ja Lu (2011) löysivät yhteyden kaikkien muiden piirteiden paitsi tunnollisuuden kanssa ja erityisesti ulospäinsuuntautuneisuuden sekä neuroottisuuden merkitys korostui. Zhang ja kumppanit (2020) taas löysivät negatiivisen yhteyden neuroottisuuden ja luottamuksen välillä.

Ulospäinsuuntautuneilla ihmisillä nähdään olevan tarve dominoida ja olla kontrollissa, jolla on selitetty esimerkiksi sitä, että he sallivat robottien lähestyä itseään silloin kun he olivat kontrollissa, verrattuna sisäänpäin suuntautuneisiin henkilöihin (Syrdal ym., 2007). Ulospäinsuuntautuneisuuden positiivinen yhteys esimerkiksi humanoideihin (eli ihmisiltä näyttävien) luottamiseen on dokumentoitu esimerkiksi Haringin ja kumppaneiden (2103) tutkimuksessa. Heidän mukaansa myös aiempien tutkimustulosten valossa näyttää siltä, että siinä missä ulospäinsuuntautuneet henkilöt ovat vuorovaikutuksessa mieluummin humanoidien robottien kanssa, sisäänpäin suuntautuneet valitsevat mieluummin ulkonäöltään mekaanisen robotin. Näin ollen tässäkin tutkimuksessa ihmisten mielikuvat ja se, minkälaiseksi robotti kuviteltiin, on saattanut vaikuttaa tutkimuksen tuloksiin. Ihmisten on myös huomattu suosivan robotteja, joilla on samankaltainen persoonallisuus ja ulospäinsuuntautuneisuus on kenties helpoin piirre ilmentää roboteissa (esimerkiksi hymyilemällä tai liioitelluilla liikkeillä) (Lionel, 2018). Ulospäinsuuntautuneisuuden tai muidenkaan persoonallisuuspiirteiden vaikutukset eivät kenties olleet vahvoja, koska tässä tutkimuksessa ei käytetty fyysistä robottia.

On myös joitain tutkimuksia, joissa ulospäinsuuntautuneisuuden tai muidenkaan piirteiden vaikutuksia ole huomattu. Esimerkiksi Salem ja kumppanit (2015) eivät löytäneet yhteyttä ulospäinsuuntautuneisuuden tai neuroottisuuden sekä luottamuksen välillä, vaikka ne olivatkin yhteydessä ulospäinsuuntautuneiden kohdalla lämpimämpään ja neuroottisuuden kohdalla kylmempään suhtautumiseen robotteja kohtaan. Samoin tekoälyä tutkittaessa on ollut myös näyttöä siitä, etteivät persoonallisuuspiirteet selittäisi luottamusta (Noah & Sethumadhvan, 2019). Aiemmat tutkimustulokset ovat siis kuitenkin ristiriitaisia ja lisää tutkimustuloksia aiheesta tarvitaan.

Hypoteesi 2 eli avoimuuden yhteys luottamukseen vahvistettiin osittain ja avoimempien ihmisten huomattiin luottavan tekoälyyn enemmän. Esimerkiksi Zhangin ja kumppaneiden (2020) tutkimuksessa, jossa mitattiin luottamusta automatisoituja ajoneuvoja kohtaan, avoimuuden ja luottamuksen välillä huomattiin yhteys. Avoimuus oli positiivisesti ja tunnollisuus negatiivisesti yhteydessä luottamukseen myös Oksasen ja kumppaneiden (2020) tutkimuksessa. Kyseisessä tutkimuksessa käytettiinkin samaa aineistoa, kuitenkin eri ryhmillä. Avoimuuden on huomattu olevan yhteydessä robottien tai tekoälyn hyväksynnän

kanssa (Conti ym., 2017). Avoimet ihmiset ovat usein uteliaita (Nickerson & Reilly, 2004) ja se on yhdistetty esimerkiksi laajempien arvojen hyväksymiseen. Tämä taas voi toimia rehellisyyden ennustajana, koska ihmiset saattavat nähdä toisten arvot samankaltaisina kuin omansa (Alarcon ym., 2018). Avoimuuden on huomattu olevan myös tärkeä tekijä siinä, minkälaisena robotti nähdään ja minkälaiseksi sen persoonallisuuden tulkitaan olevan. Avoimuus korreloi negatiivisesti roboteille annettujen ulospäinsuuntautuneisuuden ja sovinnollisuuden arvioiden kanssa Ogawan ja kumppaneiden (2018) tutkimuksessa. He olivat myös valmiimpia käyttämään rahaa kuulokkeisiin, jota robotti mainosti. Vaikka tulos johtuisikin enimmäkseen robotin suostuttelevista ominaisuuksista, voi siihen liittyä myös luottamuksen merkitys, jossa robotin mainostamaan tuotteeseen luotetaan ja näin ollen ollaan valmiita ostamaan.

Tässä tutkimuksessa löydettiin lisäksi heikko, mutta tilastollisesti merkitsevä, negatiivinen yhteys tunnollisuuden ja luottamuksen kanssa, jossa korkea tunnollisuus ennusti matalampaa luottamusta. Tunnollisuus on yhdistetty varovaisuuteen, harkitsevuuteen sekä virheiden tekemisen välttelyyn (Nickerson & Reilly, 2004). Nämä piirteet saattavat vaikuttaa luottamuspelissä siihen, että halutaan pelata varman päälle ja näin ollen vastapelaajalle annetaan vähemmän rahaa. Tunnollisuuteen liittyviä tutkimuksia etenkin luottamuksen ennustajana on kirjallisuudessa hyvin vähän. Syrdal ja kumppanit (2007) löysivät matalan tunnollisuuden omaavien henkilöiden sallivan robottien lähestyä itseään useammin tilanteissa, joissa ihminen oli kontrollissa, verrattuna tilanteisiin, joissa robotti oli kontrollissa, kun taas tällaista eroa ei huomattu korkean tunnollisuuden omaavilla. Muissa tunnollisuuteen liittyvissä tutkimuksissa esimerkiksi Cruz-Maya ja Tapus (2016) huomasivat, että tunnolliset ihmiset suoriutuivat tehtävistä paremmin, kun robotti muistutti näistä tehtävistä, verrattuna matalan tunnollisuuden omaaviin henkilöihin. Toisessa tutkimuksessa, korkean tunnollisuuden omaavien huomattiin pitävän enemmän tekstin muodossa olevasta käyttäjäliittymästä, verrattuna fyysiseen tai virtuaaliseen hahmoon, vaikka osallistujat yleisesti ottaen pitivätkin fyysistä tai virtuaalista hahmoa luotettavampana ja empaattisempana (Looije ym., 2010). Vaikuttaa siis siltä, että olisi syytä tehdä jatkotutkimuksia tähänkin persoonallisuuspiirteeseen liittyen, samalla ottaen huomioon aiemmin saadut tulokset, jotka saattavat vaikuttaa tunnollisuuden ja luottamuksen väliseen

suhteeseen.

Tässä tutkimuksessa vanhemmat vastaajat antoivat tekoälylle suurempia summia kuin nuoret. Esimerkiksi Scopelliti ja kumppanit (2005) ovat raportoineet päinvastaisesta tuloksesta, jossa nuoremmat luottivat teknologiaan enemmän kuin vanhemmat osallistujat. Iän myötä tapahtuvien kognitiivisten muutosten perusteella olisi kuitenkin loogista olettaa sen vaikuttavan luottamuksen määrään. Esimerkiksi Ho ja kumppanit (2005) löysivät, että vaikka nuoret ja vanhemmat osallistujat kalibroivat eli sopeuttivat oman luottamuksensa määrää samalla tavalla, vanhemmat henkilöt luottivat tutkimuksessa käytettyyn päätöksentekoaapuvälineeseen nuoria enemmän. He myös turvautuivat näihin välineisiin enemmän ja näin ollen tekivät enemmän laiminlyönnin sekä tehtävänannon virheitä. Tämä mitä luultavammin johtuu siitä, että automaatioon luotettaessa liikaa, sitä valvotaan vähemmän ja siihen kiinnitetään vähemmän huomiota. Sanchezin ja kumppaneiden (2004) tutkimuksessa vanhemmat osallistujat kalibroivat omaa luottamustaan nuoria paremmin, kun automaation luotettavuutta muutettiin. He olivat siis nuoria parempia oman luottamuksensa määrän muuttamisessa ja sopeuttamisessa automaation luotettavuuden muuttuessa. Tämä estää liiallisen luottamuksen tuoman itsetyytyväisyyden tai liian vähäisen luottamuksen tuoman laiminlyönnin (Lee & See, 2004). Hoffin ja Bashirin (2015) tekemän meta-analyysin perusteella iän ja luottamuksen välisen tutkimuksen tulokset ovat hyvin ristiriitaisia ja lisää tutkimuksia tarvitaan, joissa käytettäisiin esimerkiksi erilaisia, arkipäiväisiä automaatioita erilaisissa konteksteissa. Samoin enemmän kuin vain kahden ryhmän käyttö iän vaikutuksia analysoidessa voisi antaa laajemman kuvan aiheesta.

Tutkimustulosten mukaan sekä tekoälyä että robotteja kohtaan osoitettiin luottamusta, ja vaikka roboteille annettiin hieman enemmän rahaa kuin tekoälylle, ei tämä ero ollut tilastollisesti merkitsevä. Luottamuksen ilmenemistä voidaan pitää luonnollisena, koska ihmiset ovat yhä enemmän tekemisissä robottien kanssa, näkevät niitä ja kuulevat niistä esimerkiksi median kautta yhä enemmän. Samoin tekoälyn käyttö on lisääntynyt päivittäisessä käytössä. Robotteja on myös kehitetty näyttämään sekä enemmän ihmisenkaltaiselta että ihmissilmään miellyttävämmältä (Sanders ym., 2011), perustuen samankaltaisuuden sääntöön, jossa ihmiset tuntevat vetovoimaa toisia samankaltaisia henkilöitä tai asioita kohtaan (Nass & Lee, 2001).

Tarkasteltaessa aiempia tutkimuksia on tärkeää kiinnittää huomiota siihen, että tulokset saattavat vaihdella hyvinkin paljon tutkimusasetelman, kohteen sekä suoritettavan tehtävän mukaan. Myös ihmisten aikaisemmillä kokemuksilla ja taidoilla sekä muilla luottamukseen vaikuttavilla tekijöillä on suuri merkitys luottamuksen muodostumisessa, jotka tulee huomioida jatkotutkimuksissa. Persoonallisuuspiireet vaikuttavat luottamukseen eri tavalla riippuen tilanteesta tai suoritettavasta tehtävästä sekä siihen annetusta ajasta. Niiden on huomattu korreloivan positiivisesti erityisesti silloin, kun ihmiset eivät ole varmoja automaation luotettavuudesta, kun taas piirteiden merkitys vähenee ja tilannekohtaisuuden tärkeys nousee esiin, kun ihmisillä on enemmän tietoa kohteen luotettavuudesta (Szalma & Taylor, 2011).

7.2 Rajoitteet ja jatkotutkimus

Tässä tutkimuksessa ei keskitytty yhteen tiettyyn robottityyppiin, joka mahdollisti luottamuksen tutkimisen ihmisten yleisten robotti- tai tekoälykäsitysten mukaan. Koska kielellä ja sen assosiaatioilla on merkitystä ihmisten representaatioiden ja käyttäytymisen kannalta, tällä menetelmällä pystyttiin välttämään ihmisten mielikuvien laajan kirjon. Menetelmä on hyödyllinen varsinkin verkkoympäristön tutkimuksessa, jossa moni robottien tai tekoälyn ominaisuuksista ovat huomaamattomia (Savela ym., 2021). Vaikka tämä onkin yksi tutkimuksen vahvuuksista, on se samalla myös yksi sen rajoitteista. Jatkotutkimuksissa olisikin hyödyllistä tutkia luottamussuhteita myös koetilanteissa, joissa ollaan suoraan yhteydessä robotin tai tekoälyn kanssa. On tärkeää huomioida, että luottamuspelin kaltaisessa koetilanteessa ihmiset saattavat käyttäytyä eri tavalla kuin oikeassa elämässä, koska osallistujat saattavat ajatella, etteivät tule tapaamaan vastapelaajaa tulevaisuudessa ja näin ollen käyttäytymisestä ei ilmene pitkäaikaisia seurauksia (Oksanen ym., 2020).

Tutkimusten mukaan luottamus vaihtelee merkittävästi eri kulttuurien välillä. Luottamukseen vaikuttavia tekijöitä ovat muun muassa kulttuurin individualistisuus tai kollektiivisuus, kunnioitus vallalla olevia normeja kohtaan tai riippuvaisuus viranomaisiin. Esimerkiksi Karvonen (2001) tutki Pohjoismaisten kuluttajien luottamuksen ilmenemistä verkkokaupassa Suomessa, Ruotsissa ja Islannissa, jossa ilmeni, että siinä missä islantilaisilla ilmeni eniten luottamusta, suomalaiset olivat eniten varuillaan. Japanilaisten ja

yhdysvaltalaisien luottamusta tutkivassa tutkimuksessa taas yhdysvaltalaisien huomattiin luottavan automaatioon japanilaisia huomattavasti enemmän (Doney ym., 1998). Tässä tutkimuksessa maiden välinen vertailu ei ollut mahdollista, koska tutkimus toteutettiin vain yhdysvaltalaisilla vastaajilla. Näin ollen luottamukseen liittyviä tuloksia tulkittaessa ja yleistettäessä on tärkeää ottaa huomioon kulttuurinen konteksti ja varmistaa tulosten luotettavuus myös kohdemaassa sovellettaessa tuloksia muihin maihin.

Aiemmat tutkimustulokset ovat hyvinkin ristiriitaisia, joten persoonallisuuspiirteiden vaikutusta luottamukseen pitäisi tarkastella teknologian ja automaation kannalta vielä enemmän. Myös monitieteisten tutkimusten tekeminen enemmän olisi mielenkiintoista, koska robottien ja tekoälyn käytön mahdollisuus on hyvin laaja ja niiden tuomia hyötyjä on mahdollista soveltaa eri aloihin. Lionelin (2018) tekemän kirjallisuuskatsauksen perusteella, jossa robotteja kohtaan tunnetun luottamuksen ja persoonallisuuspiirteiden yhteyttä tarkasteltiin, ei löytynyt tutkimuksia, joissa esimerkiksi vuorovaikutuksen konteksti olisi otettu huomioon. Kotona ja työpaikalla olevilta roboteilta saatetaan odottaa täysin eri asioita, jonka takia jatkotutkimuksissa tähän olisi tärkeää kiinnittää huomiota. Samoin laadullisten sekä pitkittäistutkimusten puute alalla on katsauksen perusteella merkittävä.

7.3 Johtopäätökset

Persoonallisuuspiirteiden vaikutusten tutkiminen luottamuksen ilmenemiseen on tärkeää, koska tutkimusten mukaan ihmiset hyväksyvät ja tekevät helpommin yhteistyötä robottien kanssa, jotka muistuttavat käyttäjän omaa persoonallisuutta. Lisäksi on arvokasta ymmärtää, erilaisten ihmisten ja persoonien suhtautumista teknologioihin. On mahdollista oppia paljon ihmisyydestä uusien teknologian tutkimusten avulla. Näin ollen tulosten perusteella on mahdollista suunnitella esimerkiksi erilaisia robotteja, jotka mahdollistavat mahdollisimman laajan käytön sekä ihmisten hyväksynnän. Käyttämällä erilaisia sosiaalisia strategioita robotit voivat aiheuttaa ihmisillä vastavuoroisuuden reaktion, joka on esimerkiksi Kahnin ja kumppaneiden (2006) mielestä yksi ihmisten ja robottien välisen vuorovaikutuksen kulmakivistä.

Tämän tutkimuksen tavoitteena oli tutkia persoonallisuuspiirteiden yhteyttä robotteja tai

tekoälyä kohtaan tunnettuun luottamukseen. Samalla tutkittiin iän, sukupuolen, työn ja koulutuksen vaikutuksia robotteja ja tekoälyä kohtaan tunnettuun luottamukseen sekä sitä, miten luottamus robotteihin eroaa luottamuksesta tekoölyyn. Tulosten mukaan tekoälyryhmässä avoimuus ja vanhempi ikä ennustivat positiivisesti ja tunnollisuus negatiivisesti luottamusta. Robottiryhmässä ainoastaan avoimuus oli tilastollisesti merkitsevästi yhteydessä luottamuksen kanssa. Havaitut yhteydet olivat tilastollisesti merkitseviä, mutta eivät kovin voimakkaita. Osallistujien luottamus ei vaihdellut robottien ja tekoälyn välillä, ja molemmille annettiin suunnilleen sama määrä rahaa. Tässä tutkimuksessa saadut tulokset vastaavat osittain aiemmin tehtyjä tutkimuksia. Kirjallisuudessa tulokset ovat kuitenkin paikoin hyvinkin ristiriitaisia ja lisää tutkimuksia tarvitaan etenkin persoonallisuuspiirteiden ja luottamuksen välillä, jotta saataisiin kehitettyä luotettavampia ja helposti hyväksyttäviä robotteja ja tekoälyä.

Lähdeluettelo

- Alarcon, G. M., Lyons, J. B., Christensen, J. C., Bowers, M. A., Klosterman, S. L. & Capiola, A. (2018). The role of propensity to trust and the Five Factor model across the trust process. *Journal of Research in Personality*, 75, 69–82.
- Allport, G. W. & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs*, 47(1), 1–171.
- Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bartneck, C., Suzuki, T., Kanda, T. & Nomura T. (2007). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *Ai & Society*, 21, 217–230.
- Berg, J., Dickhaut, J. & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122–142.
- Bernotat, J. & Eyssel, F. A. (2017). Robot at home - How affect, technology commitment, and personality traits influence user experience in an intelligent robotics apartment. *IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)*, 26(1), 641-646.
- Bierhoff, H.-W. & Vornefeld, B. (2004). The social psychology of trust with applications in the Internet. *Analyse & Kritik*, 26, 48–62.
- Billings, D. R., Schaefer, K. E., Chen, J. Y. C. & Hancock, P. A. (2012). Human-robot interaction: developing trust in robots. *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 12(1), 109-110.
- Brühlhart, M. & Usunier, J. C. (2012). Does the trust game measure trust? *Economics Letters*, 115(1), 20–23.
- Camerer, C.F. (2003). *Behavioral game theory. Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.
- Carver, C. S. & Scheier, M. F. (2000). *Perspectives on personality* (4. painos). Boston, MA: Allyn and Bacon.
- Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I., Harrington, H. L.,... Poulton, R. (2003). Influence of life stress on depression: moderation of a polymorphism in the 5-HTT gene. *Science*, 301(5631), 386–89.

- Cattell, R. B. (1956). Second-order personality factors in the questionnaire realm. *Journal of Consulting Psychology, 20*(6), 411–418.
- Chatterjee, S., & Simonoff, J. S. (2013). *Handbook of Regression Analysis*. John Wiley & Sons, Incorporated: New Jersey.
- Chen, J. Y. C., Barnes, M. J. & Harper-Sciarini, M. (2010). Supervisory control of multiple robots: Human-performance issues and user-interface design. *IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews, 41*, 435–454.
- Cobb-Clark, D. A. & Schurer, S. (2012). The stability of big-five personality traits. *Economic Letters, 115*(1), 11–5.
- Coeckelbergh, M. (2011). Humans, animals, and robots: a phenomenological approach to human-robot relations. *International Journal of Social Robotics, 3*, 197–2014.
- Conti, D., Commodari, E. & Buono, S. (2017). Personality factors and acceptability of socially assistive robotics in teachers with and without specialized training for children with disability. *Life Span and Disability, 20*(2), 251–272.
- Costa, P. T., Jr. & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54*, 853–863.
- Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior, 46*(2), 260–281.
- Craenen, B., Deshmukh, A., Foster, M. & Vinciarelli, A. (2018). Do we really like robots that match our personality? The case of Big-Five traits, Godspeed scores and robotic gestures. *Proceedings of the International Symposium on Robot and Human Interactive Communication (RO-MAN), 27*(1), 626-631.
- Credé, M., Harms, P., Niehorster, S. & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology, 102*(4), 874.
- Cruz-Maya, A. & Tapus, A. (2017). Learning users' and personality-gender preferences in close human-robot interaction. *Robot and Human Interactive Communication (RO-MAN), 26*(1), 791–798.
- Delhey, J. & Newton, K. (2003). Who trusts? The origins of social trust in seven societies. *European Societies, 5*(2), 93–137.
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution, 2*, 265–279.
- Digman, J. M. (1990). Personality structure: Emergence of the Five Factor model. *Annual Review of Psychology, 41*(1), 417–440.

- Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly*, 34, 373–396.
- Doise, W. (1986). *Levels of Explanation in Social Psychology. European Monographs in Social Psychology*. Cambridge: Cambridge University Press.
- Doney, P. M., Cannon, J. P. & Mullen, M. R. (1998). Understanding the influence of national culture on the development of trust. *Academy of Management Review*, 23(3), 601–620.
- Driskell, J. E., Goodwill, G. F., Salas, E., & Shea, P. G. (2006). What makes a good team player? Personality and team effectiveness. *Group Dynamics: Theory, Research, and Practice*, 10(4), 249–271.
- Ermisch, J., Gambetta, D., Laurie, H., Siedler, D. & Noah Uhrig, S. C. (2009). Measuring people's trust. *Journal of the Royal Statistical Society*, 172(4), 749–769.
- Evans, A. M. & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, 42(6), 1585–1593.
- Feshbach, S., Weiner, B., & Bohart, A. (1996). *Personality* (4. painos). Lexington, MA: D.C. Heath and Company.
- Fleeson, W. & Jayawickreme, E. (2015). The whole trait theory. *Journal of Research in Personality*, 56, 82–92.
- Freedy, A., de Visser, E., Weltman, G. & Coeyman, N. (2007). Measurement of trust in human-robot collaboration. *2007 International Symposium on Collaborative Technologies and Systems*, Orlando, Florida, USA.
- Gagné, P. & Hancock, G. R. (2006). Measurement model quality, sample size, and solution propriety in confirmatory factor models. *Multivariate Behavioral Research*, 41(1), 65–83.
- George, D. & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4. painos). Boston: Allyn & Bacon.
- Gkioulekas, I. & Papageorgiou, L. G. (2019). Piecewise regression analysis through information criteria using mathematical programming. *Expert System with Applications*. 121, 362–372.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A. & Soutter, S. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115(3), 811–846.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. Teoksessa L. Wheeler (toim.) *Review of personality and social psychology* (2. painos, s. 141–165). Beverly Hills, CA: Sage.

- Goertzel, B. (2006). *The Hidden Pattern*. Brown Walker Press: Boca Raton, Florida.
- Groom, V. & Nass, C. (2007). Can robots be teammates?: Benchmarks in human–robot teams. *Interaction Studies*, 8(3), 483–500.
- Hair, J. F., Andersson, R. E., Tatham, R. L. & Black, W. C. (1998). *Multivariate Data Analysis*, (5. painos). New Jersey: Prentice Hall.
- Hakulinen, C. & Jokela, M. (2019). Persoonallisuus ja Terveys. Teoksessa: Sinikallio, S. (toim.) *Terveiden Psykologia*, (s. 49–64). PS-kustannus: Jyväskylä.
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J. & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, 53(5), 517–527.
- Haring, K. S., Matsumoto, Y. & Watanabe, K. (2013). How do people perceive and trust a lifelike robot. *Proceedings of the World Congress on Engineering and Computer Science, 1, WCECS 2013*, San Francisco, USA.
- Ho, G., Wheatley, D. & Scialfa, C. T. (2005). Age differences in trust and reliance of a medication management system. *Interacting with Computers*, 17, 690–710.
- Hoff, K. A. & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434.
- ISO 8373:22012 Robots and robotic devices – Vocabulary*. Haettu osoitteesta <https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-2:v1:en>.
- John, O. & Srivastava S. (1999): The Big Five trait taxonomy: History, measurement, and theoretical perspectives. Teoksessa Pervin, L. & John, O. (toim.) *Handbook of Personality* (s. 102–138). New York: The Guilford Press.
- Johnson, N. D. & Mislin, A. A. (2011). Trust game: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865–889.
- Kahn, P., Ishiguro, H., Friedman, B., Kanda, T. (2006). What is a human? Toward psychological benchmarks in the field of human–robot interaction. *Interaction Studies*, 8(3), 364–371.
- Karvonen, K. (2001). Designing trust for a universal audience: A multicultural study on the formation of trust in the internet in the Nordic countries. Teoksessa Stephanidis, C. (toim.) *First International Conference on Universal Access in Human-Computer Interaction*, 3, (s. 1078–1082). Mahwah, NJ: Erlbaum.
- Lahey, B. (1992). *Psychology: An Introduction*. Dubuque, IA: Wm. C. Brown Publishers.

- Lang, F. R., John, D., Lüdtke, O., Schupp, J. & Wagner, G. G. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior research methods*, 43(2), 548–567.
- Lazzarini, S. G., Madalozzo, R., Artes, R. & Siqueira, J. O. (2004). Measuring trust: An experiment in Brazil. *Inspere Working Papers*, 49, 1–27.
- Legg, S. & Hutter, M. (2007). A collection of definitions of intelligence. Teoksessa Goertzel, B. & Wang, P. (toim.) *Advances in artificial general intelligence: Concepts, Architectures and Algorithms vol 157 of Frontiers in Artificial Intelligence and Applications*. (s. 17–24). Amsterdam, NL: IOS press.
- Lewandowsky, S., Mundy, M. & Tan, G. P. (2020). The dynamics of trust: comparing humans to automation. *Journal of Experimental Psychology: Applied*, 6(2), 104–123.
- Lee, J. & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10), 1243–1270.
- Lee, J. D. & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.
- Lewicki, R. J., McAllister, D. J. & Bies, R. J. (1998). Trust and distrust: New relationship and realities. *Academy of Management Review*, 23(3), 438–458.
- Lewicki, R. J. & Bunker, B. B. (1996). Developing and maintaining trust in work relationships. Teoksessa Kramer, M. R. & Tyler, T. R. (toim.) *Trust in Organizations: Frontiers of Theory and Research* (s. 114–134). Sage Publications: London.
- Lionel, P. R. (2018 elokuu). Personality in the human robot interaction literature: a review and brief critique. *Proceedings of the 24th Americas Conference on Information Systems*, New Orleans, LA.
- Looije, R., Neerincx, M. A. & Cnossen, F. (2010). Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies*, 68(6), 386–397.
- Lucas, G. M., Gratch, J., King, A. & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94–100.
- Lönnqvist, J. E. & Tuulio-Henriksson, A. (2008). NEO-PI-R persoonallisuussmittarin suomenkielisen käännöksen validointi. *Kansanterveyslaitoksen julkaisuja*, B33/2008. Helsinki: Yliopistopaino.
- Makridakis, S. (2017). The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90, 46–60.

- Matthews, G., Deary, I. J. & Whiteman, M. C. (2003). *Personality Traits*, (2. painos). Cambridge: Cambridge University Press.
- Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734.
- McAdams, D. P. & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology*, 61, 517–542.
- McCrae, R. R. & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81–90.
- McCrae, R. R. & Costa, P. T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, 3(6), 173–175.
- McCrae, R. R., Terracciano, A. & Personality Profiles of Cultures Project (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology*, 89, 407–425.
- Meerbeek, B., Saerbeck, M. & Bartneck, C. (2009). Iterative design process for robots with personality. Teoksessa *AISB2009 Symposium on New Frontiers in Human-Robot Interaction*. *SSAISB* (s. 94–101). Springer, Berlin.
- Melo de, C. & Gratch, J. (2015). People show envy, not guilt, when making decisions with machines. Teoksessa *Proceedings of the 6th International Conference on Affective Computing and Intelligent Interaction* (s. 315-321), Xi'an, Kiina.
- Merritt, S. M., Heimbach, H., LaChapell, J. & Lee, D. (2012). I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system. *The Journal of the Human Factors and Ergonomics Society*, 55(3), 520–534.
- Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors*, 50, 194 –210.
- Meyer, J., Miller, C., Hancock, P., de Visser, E. J. & Dorneich, M. (2016). Politeness in machine-human and human-human interaction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), s. 279–283.
- Moray, N., Inagaki, T. & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied*, 6(1), 44–58.
- Morgeson, F. P., Reider, M. H. & Campion, M. A. (2005). Selecting individuals in team settings: the importance of social skills, personality characteristics, and teamwork knowledge. *Personnel Psychology*, 58, 583–611.

- Mori, M. (1970). Bukimi no tani (The uncanny valley). *Energy*, 7(4), 33–35.
- Müller, S. L. & Richert, A. (2018). The Big-Five personality dimensions and attitudes towards robots: A cross sectional study. *Proceedings of the Pervasive Technologies Related to Assistive Environments Conference*, 11, 405–408).
- Müller, J., & Schwier, C. (2019). Big Five personality factors in the Trust Game. *Journal of Business Economics*, 90, 37–55.
- Nass, C., & Lee, K. N. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology – Applied*, 7, 171–181.
- Nickerson, J. V. & Reilly, R. R. (2004, tammikuu). *A model for investigating the effects of machine autonomy on human behavior*. Esitelmä pidetty tilaisuudessa the 37th Hawaii International Conference on System Sciences, Waikoloa, Hawaii.
- Noah, B. & Sethumadhavan, A. (2019). Generational differences in trust in digital assistants. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63, 206–210.
- Nummenmaa, L. (2008). *Käyttäytymistieteiden Tilastolliset Menetelmät*. Vammala: Tammi.
- Ogawa, K., Bartneck, C., Sakamoto, D., Kanda, T., Ono, T. & Ishiguro, H. (2018). Can an android persuade you? Teoksessa Ishiguro H. & Dalla Libera F. (toim.) *Geminoid Studies* (s. 235–247). Singapore: Springer.
- Oksanen, A., Savela, N., Latikka, R., & Koivula, A. (2020). Trust Toward Robots and Artificial Intelligence: An Experimental Approach to Human–Technology Interactions Online. *Frontiers in Psychology*, 11, 568256.
- Parasuraman, R. & Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39, 230–253.
- Park, E., Jenkins, Q. & Jiang, X. (2008). *Measuring trust of human operators in new generation rescue robots*. Esitelmä pidetty tilaisuudessa the 7th JFPS International Symposium on Fluid Power, Toyama, Japani.
- Patrick, C. L. (2011). Student evaluations of teaching: effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education*, 36(2), 239–249.

- Ray, C., Mondada, F. & Siegwart, R. (2008). What do people expect from robots? *International Conference on Intelligent Robots and Systems* (s. 3816–3821).
- Richards, N. & Smart, W. (2013). How should the law think about robots? Teoksessa Calo, R. A., Froomkin, M. & Kerr, I. (toim.) *Robot Law* (s. 3–22). Cheltenham: Edward Elgar Publishing Limited.
- Robert, L., Alahmad, R., Esterwood, C., Kim, S., You, S. & Zhang, Q. (2020). A review of personality in human–robot interactions. *Foundations & Trends in Information Systems*, 4(2), 107–212.
- Robertson, I. & Cooper, G. (2011). *Well-Being: Productivity and Happiness at Work*. New York: Palgrave Macmillan.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S. & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23(3), 393–404.
- Ryckman, R. (2008). *Theories of personality* (9. painos). Belmont, CA: Thomson Wadsworth.
- Salem, M., Lakatos, G., Amriabdollahian, F. & Dautenhahn, K. (2015). Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. Teoksessa *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (s. 141–148). Portland, OR: ACM.
- Sanchez, J., Fisk, A. D. & Rogers, W. A. (2004). Reliability and age-related effects on trust and reliance of a decision support aid. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(3), 586–589.
- Sanders, T., Oleson, K. E., Billings, D. R., Chen, J. Y. C., & Hancock, P. A. (2011). A model of human-robot trust: theoretical model development. *Proceedings of The Human Factors and Ergonomics Society Annual Meeting*, 55(1), 1432–1436.
- Santamaria, T. & Nathan-Roberts, D. (2017). Personality measurement and design in human-robot interaction: a systematic and critical review. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 853–857.
- Savela, N., Kaakinen, M., Ellonen, N., & Oksanen, A. (2021). Sharing a work team with robots: The negative effect of robot co-workers on in-group identification with the work team. *Computers in Human Behavior*, 115, 106585.
- Schaefer, K. E., Chen, J. Y. C., Szalma, J. L. & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 58(3), 1–24.

- Schaefer, K. E., Sanders, T. L., Yordon, R. E., Billings, D. R. & Hancock, P. A. (2012). Classification of robot form: factors predicting perceived trustworthiness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 1548–1552.
- Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: A psychological approach. *Universal Access in the Information Society*, 4(2), 146–155.
- Seppälä, T., Olakivi, A. & Pirttilä-Backman, A.-M. (2012). Luottamus ja sosiaalipsykologisen selittämisen tasot. *Psykologia*, 47(05-06), 334–347.
- Shikishima, C., Hiraishi, K., & Ando, J. (2006). Genetic and environmental influences on general trust: A test of a theory of trust with behavioral genetic and evolutionary psychology approaches. *Japanese Journal of Social Psychology*, 22, 48–57.
- Specht, J., Egloff, B. & Schmukle, S. C. (2011) Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, 101(4), 862–82.
- Syrdal, D. S., Koay, K. L., Walters, M. L. & Dautenhahn, K. (2007). A personalized robot companion? - The role of individual differences on spatial preferences in HRI scenarios. *IEEE International Conference on Robot & Human Interactive Communication WC3-2*, 16(1), 1143–1148.
- Swope, K. J., Cadigan, J., Schmitt, P. M. & Shupp, R. (2008). Personality preferences in laboratory economics experiments. *The Journal of Socio-Economics*, 37(3), 998–1009.
- Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: The Five Factor model of personality. *Journal of Experimental Psychology: Applied*, 17, 71–96.
- Taber, K. S. (2018). The use of Cronbach’s alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, 1273–1296.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach’s alpha. *International Journal of Medical Education*, 2, 53–55.
- Tellegen A. (1985) Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. Teoksessa Tuma A. H. & Mason J. (toim.) *Anxiety and the Anxiety Disorders* (s. 681–706). Erlbaum, Hillsdale, N.J.
- Tsui, K. M. & Yanco, H. A. (2007). Assistive, rehabilitation, and surgical robots from the perspective of medical and healthcare professionals. Teoksessa *Proceedings of the*

- AAAI Workshop on Human Implications of Human-Robot Interaction* (s. 34–39). Vancouver, Canada: AAAI Press.
- Tupes, E. C. & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, 60(2), 225–251.
- Voss, P. (2005). Essentials of general intelligence: The direct path to AGI. Teoksessa Goertzel B. & Pennachin, C. (toim.) *Artificial General Intelligence* (s. 131-157). Berliini: Springer-Verlag.
- Van Lange, P. A. M. (2015). Generalized trust: Four lessons from genetics and culture. *Current Directions in Psychological Science*, 24, 71–76.
- Wang, P. (1995). *On the working definition of intelligence*. Technical Report 94, Center for Research on Concepts and Cognition, Indiana University.
- Weiss, A., Bernhaupt, R. & Lankes, M. (2008). The influence of control on the acceptance of ambient intelligence by elderly people: An explorative study. Teoksessa Aarts, E. H. L., Crowley, J. L., Ruyter, B. E. R., de Gerhauser, H., Pflaum, A., Schmidt, J. ym. (toim.) *Ambient Intelligence - Europeanconference* (s. 58–74). Nuremberg, Saksa.
- Wood, S. (2012). Prone to progress: Using personality to identify supporters of innovative social entrepreneurship. *Journal of Public Policy and Marketing*, 31(1), 129–141.
- Worchel, P. 1979. Trust and distrust. Teoksessa W. G. Austin & S. Worchel (toim.) *The social psychology of intergroup relations*. Belmont, CA: Wadsworth.
- Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H. & Zhu, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies*, 112, 220–223.
- Zhou, T. & Lu, Y. (2011). The effects of personality traits on user acceptance of mobile commerce. *International Journal of Human-Computer Interaction*, 27(6), 545–561.

Liitteet

Liite A.

Regressiomallien vif-arvot robottiryhmässä

Muuttuja	VIF	1/VIF
Neuroottisuus	1,40	0,72
Tunnollisuus	1,36	0,73
Sovinnollisuus	1,25	0,80
Avoimuus	1,21	0,83
Ulospäin- suuntautuneisuus	1,20	0,83
VIF Keskiarvo	1,29	

Liite B.

Regressiomallien vif-arvot tekoälyryhmässä

Muuttuja	VIF	1/VIF
Neuroottisuus	1,23	0,81
Tunnollisuus	1,41	0,71
Sovinnollisuus	1,26	0,79
Avoimuus	1,16	0,87
Ulospäin- suuntautuneisuus	1,11	0,90
VIF keskiarvo	1.23	