

Juuso-Pekka Siivonen

**CHATBOTTIEN SYNNYTTÄMIEN SOSIAALISTEN
REAKTIOIDEN YKSILÖLLISYYS**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2021

TIIVISTELMÄ

Siivonen, Juuso-Pekka

Chatbottien synnyttämien sosiaalisten reaktioiden yksilöllisyys

Jyväskylä: Jyväskylän yliopisto, 2021, 81 s.

Tietojärjestelmätiede, pro gradu -tutkielma

Ohjaaja: Kyppö, Jorma

Tietotekniikan kehityksen ansiosta erilaiset chatbot-toteutukset ovat yleistyneet nykyisessä informaatioteknisessä ympäristössä ja niitä hyödynnetään niin viih-teellisiin kuin toimintaa tehostaviin tarkoituksiin. Chatbottien käytön suhteelli-sen nopeasta kasvusta huolimatta, niiden käyttökokemus ja toiminnallisuus eivät kuitenkaan yleensä yllä vastaamaan käyttäjien odotuksia. Tietokoneita on käytetty jo useamman vuosikymmenen ajan, jonka takia niiden käyttö on tuttua käyttäjille sekä aiheen tutkimukselle. Chatbottien käyttö eroaa kuitenkin suu-resti näistä opituista konventioista, sillä suorien komentojen ja kuvakkeiden klikkailun sijaan, chatbotin kanssa vuorovaikutetaan keskustelunomaisen dia-login välityksellä. Tämä keskustelunomainen toiminta tuo vuorovaikutukseen uuden sosiaalisen ulottuvuuden, jota ihmiset ovat harjoittaneet vuosituhansien ajan viestiessään toisten ihmisten kanssa. Keskustelu chatbotin kanssa saa ihmi-set tulkitsemaan ja arvioimaan chatbottia samanlailla kuin jos keskustelu-kumppanina olisi toinen ihminen. Chatbottien nähdään sisältävän sosiaalisiksi vinkeiksi kutsuttuja ominaisuuksia, joita tulkitsemalla ihmisten sosiaaliset reak-tiot chatboteista syntyvät. Chatbottien paremman suunnittelun ja toteutuksen saavuttamiseksi, ihmisen ja chatbotin välisen vuorovaikutuksen erityispiirtei-den ymmärtäminen on olennaista. Tämän tutkimuksen tavoitteena oli tuottaa syvempää tietämystä chatbottien synnyttämien sosiaalisten reaktioiden muo-dostumisesta. Tutkimus toteutettiin määrällisenä verkkokyselynä, jossa neljän eri chatbotin synnyttämiä sosiaalisia reaktioita mitattiin kolmen eri ulottuvuu-den kautta: antropomorfismin, miellyttävyyden sekä havaitun älykkyyden. Näiden ulottuvuuksien saamia arvoja vertailtiin eri chatbottien kohdalla, jonka perusteella tunnistettiin käyttäjien sosiaalisten reaktioiden eroja. Tutkimuksen tuloksina havaittiin, että chatbottien synnyttämät sosiaalisissa reaktioissa ole-van tilastollisesti merkitseviä eroja käyttäjien välillä. Lisäksi tuloksissa oli viit-teitä siitä, että chatbotin sisältämien sosiaalisten vinkkien määrän kasvaessa, myös niiden synnyttämät sosiaaliset reaktiot ovat voimakkaampia. Enemmän sosiaalisia vinkkejä sisältävien chatbottien synnyttämässä sosiaalisissa reaktiois-sa havaittiin myös vähemmän hajontaa käyttäjien välillä, verrattuna vähemmän sosiaalisia vinkkejä sisältävien chatbottien synnyttämiin sosiaalisiin reaktioihin.

Asiasanat: chatbot, ihmisen ja tietokoneen välinen vuorovaikutus, sosiaalinen reaktio, antropomorfismi, käyttökokemus, sosiaalinen vinkki

ABSTRACT

Siivonen, Juuso-Pekka

The individuality of the social reactions generated by chatbots

Jyväskylä: University of Jyväskylä, 2021, 81 pp.

Information Systems, Master's Thesis

Supervisor: Kyppö, Jorma

Thanks to the development of information technology, various chatbot implementations have become more common in the current information technology environment and are utilized for both entertainment and operational efficiency purposes. However, despite the relatively rapid growth in the use of chatbots, their user experience and functionality generally do not meet users' expectations. Computers have been used for several decades, which is why their use is familiar to users and to the research of the subject. However, the use of chatbots differs greatly from these formerly learned conventions, as instead of clicking on icons, the chatbot is interacted through a conversational dialogue. This conversational activity brings to the interaction a new social dimension that people have been using for ages when communicating with other people. Conversing with a chatbot makes people interpret and rate the chatbot in the same way as if another person were the interlocutor. Chatbots are seen to contain features called social cues, which, when interpreted, creates social reactions in people towards chatbots. In order to achieve better design and implementation of chatbots, an understanding of the specificities of human-chatbot interaction is essential. The aim of this study was to provide a deeper understanding about the formation of social reactions generated by chatbots. The study was conducted as a quantitative online survey in which the social responses generated by four different chatbots were measured through three different dimensions: anthropomorphism, pleasantness, and perceived intelligence. The values obtained for these dimensions were compared between the different chatbots, based on which differences in users' social reactions were identified. The results of the study found that there were statistically significant differences in social responses generated by chatbots between users. In addition, there were indications in the results that as the number of social cues contained in the chatbot increases, so do the social reactions they generate. Social reactions generated by chatbots with more social cues also showed less dispersion among users compared to social reactions generated by chat bots with less social tips.

Keywords: chatbot, human-computer interaction, social reaction, anthropomorphism, user experience, social cue

KUVIOT

KUVIO 1 Tyypillinen chatbotin arkkitehtuuri	14
KUVIO 2 Morin outo laakso	22

TAULUKOT

TAULUKKO 1 Tietokoneen sosiaalista läsnäoloa herättävät sosiaaliset vinkit .	20
TAULUKKO 2 Chatbotin sosiaalisten vinkkien luokittelu	26
TAULUKKO 3 Tutkimuksen chatboteista tunnistetut sosiaaliset vinkit.	38
TAULUKKO 4 Mukautettu ja suomennettu Godspeed-kysely	40
TAULUKKO 5 Kyselyn vastaajien demografiset tiedot	44
TAULUKKO 6 Mukautetun Godspeed-kyselyn summamuuttujien reliabiliteetti	45
TAULUKKO 7 Chatbottien summamuuttujien tunnusluvut	45
TAULUKKO 8 Levenen testin tulokset	46
TAULUKKO 9 Kruskal-Wallis-testin tulokset antropomorfismin ja eri chatbottien välillä	46
TAULUKKO 10 Parivertailun tulokset antropomorfismin ja chatbottien välillä	46
TAULUKKO 11 Miellyttävyyden ja havaitun älykkyyden yksisuuntaisten varianssianalyysien tulokset eri chatbottien suhteen.....	47
TAULUKKO 12 Varianssianalyysien parivertailun tulokset miellyttävyydelle ja havaitulle älykkyydelle chatbottien suhteen.....	48
TAULUKKO 13 Demografisten tekijöiden tilastolliset merkitsevyydet summamuuttujien suhteen eri chatboteille	49
TAULUKKO 14 Chatbottien summamuuttujien tilastollisesti merkitsevät erot ikäryhmien välillä.....	50
TAULUKKO 15 Chatbottien summamuuttujien tilastollisesti merkitsevät erot sukupuolien välillä	50
TAULUKKO 16 Avoimiin kysymyksiin saatujen vastausten määrä.....	51
TAULUKKO 17 Tilastollisesti merkitsevästi erot chatbottien välillä antropomorfismin, miellyttävyyden ja havaitun älykkyyden keskiarvojen suhteen	56

SISÄLLYS

TIIVISTELMÄ	2
ABSTRACT	3
KUVIOT	4
TAULUKOT	4
SISÄLLYS.....	5
1 JOHDANTO.....	7
2 KOMENNOISTA VUOROPUHELUKSI	10
2.1 Chatbot.....	10
2.2 Chatbottien historia ja kehittyminen	11
2.3 Nykyisten chatbottien yleispiirteet	13
3 IHMISEN JA TIETOKONEEN VÄLINEN VUOROVAIKUTUS.....	16
3.1 Ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimus	16
3.2 Vuorovaikutus chatbotin kanssa	17
3.3 Tietokoneet ovat sosiaalisia toimijoita.....	19
4 CHATBOTTIEN SYNNYTTÄMÄT SOSIAALISET REAKTIOT KÄYTTÄJISSÄ	23
4.1 Ihmisen ja chatbotin väliseltä vuorovaikutukselta odotetut sosiaaliset reaktiot.....	23
4.2 Chatboteissa ilmenevät sosiaaliset vinkit.....	24
4.2.1 Verbaaliset vinkit.....	26
4.2.2 Visuaaliset vinkit	28
4.2.3 Auditiiiviset vinkit	30
4.2.4 Näkymättömät vinkit	30
4.3 Sosiaalisten vinkkien tulkintaan vaikuttavia tekijöitä	31
5 TUTKIMUSMENETELMÄ JA TUTKIMUKSEN TOTEUTUS.....	34
5.1 Tutkimuksen tavoite	34
5.2 Tutkimusmenetelmän valinta	35
5.3 Tutkimuksen toteutus	36
5.3.1 Chatbot-keskustelutilanteiden luonti.....	36
5.3.2 Sosiaalisten reaktioiden mittaaminen	38
5.3.3 Kyselyn toteutus.....	40
5.4 Aineiston analysointi.....	41

6	TUTKIMUKSEN TULOKSET	43
6.1	Aineiston kuvailu	43
6.2	Mukautetun Godspeed-kyselyn tulokset.....	44
6.3	Avointen kysymyksien tulokset	50
6.3.1	Eliza	51
6.3.2	Alice.....	51
6.3.3	Mitsuku.....	52
6.3.4	EvieBot	53
7	POHDINTA	54
7.1	Tulosten tulkinta	54
7.2	Tutkimuksen arviointi ja jatkotutkimusaiheet	59
8	YHTEENVETO	61
	LÄHTEET	63
	LIITE 1 SUMMAMUUTTUJIEN HISTOGRAMMIT CHATBOTEITTAIN	71
	LIITE 2 VERKKOKYSELYLOMAKE	72

1 JOHDANTO

Erilaiset chatbotit ovat nykypäivänä tuttu näky eri nettipalvelujen sivuilla sekä mobiililaitteiden sovelluksissa. Chatbottien käyttö on kasvanut viime vuosina huomasti ja kasvun ei nähdä hiipuvan vielä tulevina vuosinakaan. Niitä käytetään niin viihteellisiin tarkoituksiin, kuin myös toimintaa tehostavissa sovelluksissa. Chatboteilla tarkoitetaan yleisesti tietokoneohjelmia, joiden kanssa ihminen vuorovaikuttaa keskustelua muistuttavan dialogin kautta (McTear, 2017). Tekoälyn kehityksen ansiosta, tämä vuorovaikutus chatbottien kanssa voi tapahtua niin tekstin välityksellä viestejä kirjoittamalla, kuin äänen välityksellä chatbotin kanssa oikeasti puhumalla. Tämänlainen vuorovaikutus eroaa paljon totutusta hiiren ja näppäimistön käytöstä, mahdollistaen siten uudenlaisia mahdollisuuksia tietokoneiden käytölle, mutta luoden sille samalla myös uusia haasteita.

Ihmisen ja chatbotin välinen vuorovaikutus on suhteellisen tuore ja vähän tutkittu tutkimusalue, jonka takia chatbotin kanssa vuorovaikuttaminen ei ole ongelmatonta. Aiheen vähäisen tutkimuksen takia, chatbottien suunnitteluun ja toteutukseen ei ole vielä muodostunut vakiintuneita käytänteitä tai standardeja, jonka takia niiden toteutuksissa nojaututaan tutkitun tiedon sijaan enemmänkin intuitioon (McTear, 2018). Tämä heijastuu myös chatbottien käyttökokemukseen, jossa usein ongelmana on chatbottien toiminnallisuuden heikko välittyminen käyttäjille (Luger & Sellen, 2016; Hill, Ford & Farreras, 2015; Zamora, 2017; Chen & Wang, 2018). Ihmisen ja chatbotin välisen vuorovaikutuksen tutkimuksella on siten olennainen osa chatbottien käyttökokemuksen parantamisessa tulevaisuudessa, jonka takia aiheen tutkiminen on tärkeää.

Ihmisen ja chatbotin välisen vuorovaikutuksen selittämistä hankaloittaa kuitenkin myös vuorovaikutuksen sosiaalinen ulottuvuus. Chatbottien kanssa käytävä keskustelunomainen vuorovaikutus muistuttaa ihmisten välistä vuorovaikutusta, jonka takia ihmisille on ominaista tulkita sitä samaan tapaan kuin ihmisten välistä vuorovaikutusta (McTear, 2017; Gnewuch, Morana & Maedche, 2017). Tämä näkyy esimerkiksi inhimillisten ominaisuuksien liittämisenä chatbotteihin, esimerkiksi chatbotin pitämisenä älykkäänä tai tulkitsemalla sen omaavan persoonallisuuden. Chatbotteihin kohdistuukin paljon ennakko-

odotuksia sen älykkyyteen ja inhimillisyyteen liittyen, joiden täytyminen jääkin usein vajavaiseksi (Luger & Sellen, 2016; Zamora, 2017; Portela & Granell-Canut, 2017; Jain, Kumar, Kota & Patel, 2018).

Tietokoneista syntyvien sosiaalisten reaktioiden nähdään syntyvän tietokoneissa olevista niin sanotuista sosiaalisista vinkeistä, joita ihmiset tulkitsevat ja joiden perusteella sosiaaliset reaktiot muodostuvat (Nass & Moon, 2000; Fogg, 2002). Myös chatbottien on havaittu sisältävän tällaisia sosiaalisia vinkkejä (Feine, Gnewuch, Morana & Maedche, 2019). Chatbottien sisältämien sosiaalisten vinkkien tutkimus ei kuitenkaan vielä ole hirveän pitkällä, vaikka niitä onkin tunnistettu ja eritelty. Aiemmassa kirjallisuudessa ihmisen ja chatbotin välisen dialogin interaktiivisuuden on havaittu kasvattavan chatbotin koettua ihmismäisyyttä (Go & Sundar, 2019; Portela & Granell-Canut, 2017; Jain ym., 2018). Samaten chatbotin visualisoinnin realistisuuden on tutkittu voimistavan chatbotin synnyttämiä sosiaalisia reaktioita (Gong, 2008; Go & Sundar, 2019; Feine ym. 2019; Araujo, 2018; Rincón-Nigro & Deng, 2013; Beun, De Vos & Witteman, 2003; Ciechanowski, Przegalinska, Magnuski & Gloor, 2018). Aiempaa tutkimusta ei olla kuitenkaan tehty chatbotin sisältämien sosiaalisten vinkkien näkökulmasta, vaikka sosiaalisilla vinkeillä vankka teoreettinen perusta sosiaalisten reaktioiden synnyssä.

Tämän tutkimuksen tavoitteena on rakentaa syvempää tietoa Feinen ym. (2019) teorian pohjalta ja selvittää tarkemmin, miten chatbottien sisältämät sosiaaliset vinkit vaikuttavat chatboteista syntyviin sosiaalisiin reaktioihin. Tutkimuksessa pyritään löytämään vastaukset seuraaviin tutkimuskysymyksiin:

- Ovatko chatbottien synnyttämät sosiaaliset reaktiot yhteneviä käyttäjien välillä?
- Vaikuttaako chatbottien sisältämien sosiaalisten vinkkien määrä chatbottien synnyttämiin sosiaalisiin reaktioihin käyttäjissä?

Tutkimusongelmien selvittämiseksi toteutettiin määrällinen tutkimus, jonka aineisto kerättiin verkkokyselyn avulla. Verkkokyselyssä mitattiin ihmisille chatboteista syntyviä sosiaalisia reaktioita kolmen eri ulottuvuuden kautta, joita olivat chatbotin koettu antropomorfismi, miellyttävyys, sekä havaittu älykkyys. Näiden kolmen ulottuvuuden saamia arvoja vertailtiin eri chatbottien kesken, joiden perusteella saaduista tuloksista tulkittiin chatbottien synnyttämien sosiaalisten reaktioiden muodostumista.

Tutkimuksen tuloksista havaittiin, että chatbottien synnyttämät sosiaaliset reaktiot eivät ole yhteneviä käyttäjien välillä, vaan niissä on eroa käyttäjien välillä. Lisäksi tuloksissa oli viitteitä siitä, että sosiaalisten vinkkien määrän kasvaessa, myös chatbotista syntyvien sosiaalisten reaktioiden voimakkuus kasvaa. Chatbotin sisältämien sosiaalisten vinkkien määrän havaittiin myös vähentävän käyttäjien välisiä eroja chatboteista syntyneissä sosiaalisissa reaktioissa.

Tutkielman rakenne on seuraava: Johdannon jälkeen, luvuissa 2, 3 ja 4 käydään läpi tutkimuksen teoreettista taustaa ja aiempaa kirjallisuutta. Toisessa luvussa esitellään chatbottien käsite yleisesti sekä miten tekoälyn kehittyminen

on johtanut chatbottien kehitystä tähän päivään saakka. Lisäksi esitellään nykyisten chatbottien yleispiirteitä. Kolmannessa luvussa siirrytään tarkastelemaan ihmisen ja tietokoneen välistä vuorovaikutusta tieteenalana, sekä käydään läpi sen ympärillä esiintyviä yleisimpiä käsitteitä. Tämän jälkeen ihmisen ja tietokoneen välistä vuorovaikutusta tarkastellaan chatbottien käytön näkökulmasta ja nostetaan esille vuorovaikutuksen sosiaalinen ulottuvuus. Jatkuksena tälle, luvun lopussa käsitellään tietokoneiden synnyttämiä sosiaalisia reaktioita ihmisissä, sekä miten ihmiset yhdistävät inhimillisiä ominaisuuksia tietokoneisiin. Neljännessä luvussa näitä ihmisen kokemia sosiaalisia reaktioita tarkastellaan rajatummin chatbotteihin liittyen. Luvussa käydään läpi ihmisten chatbotteihin kohdistamia ennakko-odotuksia, chatboteissa ilmeneviä sosiaalisia vinkkejä ja niiden kategorioita, sekä näiden sosiaalisten vinkkien tulkintaan vaikuttavia seikkoja.

Luvuissa 5, 6, 7 ja 8 käsitellään tutkimuksen empiiristä osuutta ja tuloksia. Viidennessä luvussa käydään läpi tutkimuksen tutkimusmenetelmä, sekä tutkimuksen toteutus ja niissä tehtyjen valintojen perustelut. Kuudennessa luvussa voidaan aineistonkeruussa saatuja tuloksia, joita analysoimalla muodostetaan tutkimuksen tulokset. Seitsemännessä luvussa tutkimuksen tuloksia analysoidaan ja niiden merkitystä arvioidaan. Lisäksi käydään läpi tutkimuksen rajoitteet ja jatkotutkimusaiheet. Tutkielman lopuksi on yhteenveto, käytetyt lähteet sekä viimeisenä tutkielman liitteet.

2 KOMENNOISTA VUOROPUHELUKSI

Tässä luvussa esitellään chatbotteja yleisellä tasolla ja ilmiönä. Aluksi käsitellään chatbotin käsitettä ja sen määritelmää. Tämän jälkeen chatbottien taustaa avataan kuvaamalla tiivistetysti tekoälyn kehityksen historiaa ja sen merkityksellisiä edistysaskelia chatbottien kehittymisen suhteen. Luvun lopussa nostetaan esille huomioita nykyisille chatboteille ominaisista yleispiirteistä ja käytännöistä.

2.1 Chatbot

Chatbot-sovellutukset ovat yleistyneet huimasti viime vuosina. Aikoinaan niihin saattoi törmätä esimerkiksi televisiossa näkyneenä puhuvana päänä tai Cleverbotin (Cleverbot, 2020) kaltaisten nettisivuilla viihdyttäneiden keskustelukaverien kautta. Nykyään kuitenkin chatbotit ovat levinneet laajalti ja nettisivujen alareunaan pongahtavat asiakaspalvelevat chatbotit ovat tuttu näky useimmilla nettisivuilla. Chatbotit ovat muuttuneet hetken viihdearvoa tarjoavista sovelluksista olennaisiksi osiksi eri palvelujen toimintaa. Kiinnostus chatbotteihin on kasvanut räjähdysmäisesti ja niiden sovelluksilla nähdään olevan potentiaalia varsinkin opetuksen, terveydenhuollon ja myynnin toimialoilla (Følstad & Brandtzæg, 2017). Chatbotit ovat kykeneviä muodostamaan käyttäjilleen miellyttävän ja laadukkaan vuorovaikutuskokemuksen, joka selittää niiden suotien kasvua verkkopalveluissa (Chung, Ko, Joung & Jin Kim, 2018) Gartner ennustaa 70 % toimistotyöntekijöistä vuorovaikuttavan chatbottien kanssa päivittäin vuoteen 2022 mennessä (Goasduff, 2019). Useat palvelut tarjoavat chatbotteja verkkopalvelujen sijaan myös suoraan mobiililaitteen viestintäsovelluksiin, kuten esimerkiksi Facebook Messengeriin (Araujo, 2018). Googlen, Amazonin ja Applen kaltaiset teknologiajätit ovat kehittäneet omat virtuaaliavustajina tunnetut chatbottinsa, joiden avulla älylaitteiden ohjaaminen onnistuu helposti pelkän puheen välityksellä.

Mutta mitä nämä paljon puhutut chatbotit oikeastaan ovat? McTear (2017) määrittelee chatbotin tiettyjä toimintoja suorittavaksi tietokoneohjelmistoksi, jonka kanssa käyttäjä on vuorovaikutuksessa dialogin välityksellä. Hänen mukaansa chatbotit voidaan jaotella tehtäväpainotteisiin ja keskustelupainotteisiin chatbotteihin. Erona on, että tehtäväpainotteisten chatbottien tavoitteena on suorittaa jokin toiminto, kun taas keskustelupainotteisiin chatbotteihin, joiden tarkoitus on enemmän viihteellinen ja tarjota mahdollisimman ihmismäinen keskustelukokemus (McTear, 2017). Chatbot-sovellutuksista käytetään kuitenkin monia eri nimityksiä, kuten keskusteleva käyttöliittymä, keskustelubotti, virtuaaliavustaja ja keskusteleva agentti. Nimityksillä on tapana vaihdella eri lähteiden välillä ja eri termejä käytetään yleensä kuvaamaan eri toiminnallisuu- den omaavia chatbotteja, mutta niillekään ei ole olemassa selkeitä määritelmiä. Tässä tutkielmassa käytetään termiä "chatbot" kuvaamaan yleisesti tietokoneohjelmia, joiden kanssa käyttäjä on vuorovaikutuksessa luonnollista kieltä käyttäen.

2.2 Chatbottien historia ja kehittyminen

Chatbottien taustalla on tekoälyn tutkimuksen nousu ja sen mahdollistamat teknologian kehitysaskleet viime vuosikymmenten aikana. Tietokoneen älykkyys on kiinnostanut ihmisiä jo ensimmäisten tietokoneiden syntyajoista lähtien. Yksi ensimmäisistä aiheen teorioista on McCullochin ja Pittsin (1943) esittämä matemaattinen malli ihmisaivoja mukailevasta neuroverkosta, jota pidetäänkin nykyisen tekoälyajattelun perustana. Brittiläinen matemaatikko Alan Turing (1950) kehitti 50-luvulla tunnetun Turingin testin tietokoneen älykkyyden mittaamiseen. Testin mukaan tietokone on älykäs, mikäli ihminen ei osaa erottaa keskusteleeko ihmisen vai tietokoneen kanssa. Goertzelin ja Pennanchinin (2007) mukaan Turingin testiä ei kuitenkaan nykyisin käytetä tietokoneen älykkyyden mittaamiseen, vaan tekoäly jaetaan vahvaan ja heikkoon tekoälyyn. Vahvalla tekoälyllä tarkoitetaan ihmisen kaltaisen tai kehittyneemmän tietoisuuden omaavaa tekoälyä, joka on vielä nykypäivänä täysin teoreettinen ajatus. Heikko tekoäly pitää sisällään nykyiset tekoäly- ja koneoppisovellutukset, jotka kykenevät älykkäisiin, mutta rajattuihin tehtäviin (Goertzel & Pennanchin, 2007).

Tekoälyn sovellutusten kehitysaskleet olivat pitkän aikaa pelkän tietokoneiden laskentatehon kasvun ansiota. Vuoden 1997 IBM:n Deep Blue -tietokoneen voittoa shakin hallitsevaa maailmanmestaria vastaan pidetään yhtenä tekoälyn merkittävistä merkkipaaluista, joskin voitto oli tietokoneen älykkyyden sijaan puhtaan laskentatehon ansiota: Deep Blue pystyi simuloimaan 200 miljoona pelisiirtoa sekunnissa, jonka ansiosta se pystyi valitsemaan otollisimman siirron pelitilanteen mukaan (Campbell, Hoane & Hsu, 2002). Sen sijaan Googlen AlphaGo-ohjelmiston voitto ihmistä vastaan go-pelissä todisti tietokoneen kyvykkyyttä päätöksenteon lisäksi älykkääseen toimintaan. Pelkkä laskentateho ei mitenkään riitä go-pelin kaikkien mahdollisten siirtojen arviointiin, joten AlphaGo hyödynsi neuroverkkoteknologioita ja syväoppimista

parhaan siirron tunnistamiseen ja suorittamiseen (Wang ym., 2016). Neuroverkkojen ja syväopin hyödyntäminen on mahdollistanut nykyisen tekoälysovellutusten räjähdysmäisen kasvun. Teknologijatit investoivatkin tekoälyn tutkimukseen ja kehitykseen 20–30 miljardia dollaria jo pelkästään vuonna 2016 (Chui, 2017).

Chatbottien kehitys kulkee vahvasti käsi kädessä luonnollisen kielen prosessoinnin kehityksen edistysaskeleiden kanssa. Luonnollisen kielen prosessoinnilla tarkoitetaan tekstin analysointiin ja esittämiseen käytettyjä teorioita ja teknologioita, joiden tavoitteena on saavuttaa ihmismielen tasoinen kielen prosessointi ja sen semantiikan ymmärtäminen eri käyttötarkoituksia varten (Liddy, 2001). Luonnollisen kielen prosessointi on kehittynyt sääntöpohjaisista, avainsanoja tunnistavista ohjelmistoista neuroverkkopohjaisiin toteutuksiin, jotka kykenevät tunnistamaan, luomaan ja ymmärtämään puheen ja tekstin semantiikkaa ja siinä esiintyviä asiayhteyksiä (Hirschberg & Manning, 2015). Itse puheen ja tekstin ymmärtämisen lisäksi, internetin ja massadatan hyödyntäminen mahdollistaa kontekstiin liittyvän tiedon hankkimiseen jopa reaaliajassa, jonka avulla chatbotit pystyvät suoriutumaan entistä älykkäämmistä tehtävistä ja toiminnoista (McTear, 2017).

Vaikka chatbottien yleistyminen on ollut nopeaa viime vuosien aikana teknologian kehityksen ansiosta, chatbotit eivät kuitenkaan ole ilmiönä uusi. Yksi varhaisimmista chatboteista on Joseph Weizenbaumin (1966) 60-luvulla kehittämä Eliza, jonka hän loi imitoimaan psykoterapeuttia. Eliza toimi noudattamalla ennalta määriteltäviä sääntöjä, joiden mukaisesti se etsi keskustelukumppanin viesteistä avainsanoja ja esitti niihin sopivia kommentteja tai jatkokysymyksiä (Weizenbaum, 1966). Eliza onnistui aikoinaan ällistyttämään ihmisiä luonnollisen kuuloisilla vastauksillaan, vaikkakin sen toiminta perustui kuitenkin hyvin rajattujen ohjeiden seuraamiseen, eikä siten kyennyt oikeasti älykkäiden toimintojen toteuttamiseen.

Chatbotteja voidaanakin jakaa eri kategorioihin niiden toimintatavan perusteella. McTear (2002) jakaa chatbotit kolmeen kategoriaan: Tilaperusteiset (engl. fixed-state based) järjestelmät perustuvat tiukasti ennalta määrätyn ohjeen noudattamiseen ja ovat toiminnaltaan kysymys-vastaus tyyppisiä. Kehysperusteiset (engl. frame-based) järjestelmät osaavat tulkita käyttäjän viestejä tietyn kehyksen sisällä ja eritellä syötteestä kontekstiin liittyvää tietoa. Agenttiperusteiset (engl. agent-based) järjestelmät ovat edistynein tyyppi näistä kolmesta. Niissä chatbot toimii tiukkojen ohjeiden sijaan itsenäisenä toimijana ja kykenee hahmottamaan kontekstia laajemmin, siten toimien myös monimutkaisemmissa tilanteissa. (McTear, 2002).

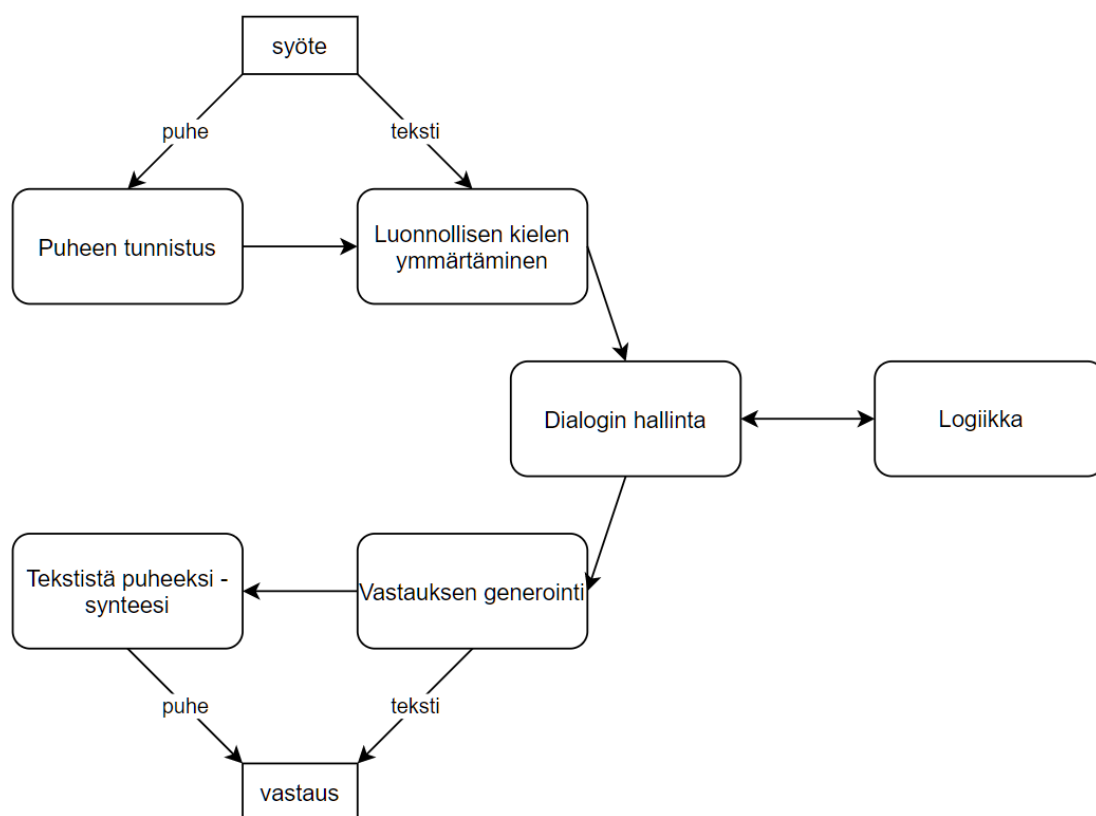
Chatbottien kasvaneen suosion taustalla on niiden kehittämiseen tarvittavien työkalujen saatavuuden helpottuminen. McTear (2017) selittää chatbottien laajan leviämisen olevan niiden kehitystyökalujen kehittymisen ja chatbottien helpottuneen integroimisen ansiota. Varhaiset chatbotit olivat hyvin herkkiä käyttäjän antamille väärille syötteille, sillä ne olivat usein kehitetty varta vasten tiettyä käyttöympäristöä ja -tehtävää varten. Tämän takia chatbotit olivat toiminnaltaan ja rakenteeltaan epävarmoja, jonka takia niiden skaalaus ja siirtämi-

nen eri käyttöympäristöön oli haasteellista. Nykyisin chatbottien kehitykseen on monia eri toimijoiden kehittämää rajapintoja ja teknologioita. Chatbottien käyttöönotto on nykyään myös huomattavasti helpompaa, sillä esimerkiksi nettisivulla toimiva chatbot toimii suoraan palvelimella, eikä vaadi käyttäjältä erillisiä toimenpiteitä (McTear, 2017).

Nykyisin tunnetuimpia chatbotteja ovat Googlen, Applen, Amazonin ja muiden teknologiajättien kehittämät chatbotit. Nämä chatbotit ovat nousseet uutisotsikoihin niiden kehittyneiden ominaisuuksien ansiosta. Esimerkiksi Google Assistant kykenee vuorovaikutteiseen keskusteluun, ohjailemaan muita tietoteknisiä laitteita ja jopa soittamaan ravintolan puhelinnumeroon tehdäksään pöytävarauksen puhumalla ravintolan asiakaspalvelijan kanssa (Google, 2020). Keskustelutaidoiltaan yksi tämän hetken kehittyneimmistä chatboteista on Steve Worswickin kehittämä Mitsuku-chatbot, joka on voittanut Loebner-palkinnon yhteensä viisi kertaa, joista neljä putkeen vuosina 2016–2019 (Mitsuku, 2021). Loebner-palkinto on vuosittainen kilpailu, jossa palkitaan ihmismäisimpiä chatbotteja. Google on myös paljastanut kehittävänsä uutta Meenanimestä chatbottia, joka heidän omista testeistään on eniten ihmiseen verrattavissa nykyisistä chatboteista, ohittaen esimerkiksi edellä mainitun Mitsukun (Adiwardana ym., 2020).

2.3 Nykyisten chatbottien yleispiirteet

Nykyisten chatbottien yleisarkkitehtuuri on säilynyt suhteellisen muuttumattomana sen taustalla olevien teknologisten toteutuksien kehittymisestä huolimatta. McTearin (2018) esittämä malli chatbotin tyypillisestä arkkitehtuurista näkyy kuviossa 1. Mallin mukaan käyttäjä antaa syötteen joko puheella tai tekstillä, jonka sisältö analysoidaan luonnollisen kielen prosessointia hyödyntämällä. Tämän jälkeen järjestelmä suorittaa toimintalogiikkansa mukaisen toiminnon syötteen perusteella. Lopuksi järjestelmän vastaus generoidaan luonnolliselle kielelle ja esitetään käyttäjälle tekstinä tai puheena (McTear, 2018).



KUVIO 1 Tyypillinen chatbotin arkkitehtuuri (McTear, 2018)

McTearin (2017) mukaan chatbottien keskusteluun perustuva toimintaperiaate ohjaa niiden vuorovaikutusta, toteutusta ja käyttöä. Tämä ilmenee chatboteissa kahdella eri tavalla: Ensimmäinen tapa liittyy chatbottien vuorovaikutukseen. Kuten aikaisemmin mainittiin, chatbottien kanssa vuorovaikutus tapahtuu luonnollisen kielen välityksellä keskustelemalla. Tämän takia dialogi chatbotin kanssa ei ole ennalta määriteltyä ja saman asian suorittaminen voi tapahtua hyvinkin erilaisia ilmaisuja käyttäen. Tehtävien ilmaisemisen vapaus lisää järjestelmien joustavuutta eri käyttötilanteisiin. Toinen chatbottien keskustelunomaisuuden tuoma ilmiö on chatbottien käyttöliittymien viesteihin perustuva rakenne (McTear, 2017). Useimmat chatbotit muistuttavat käyttöliittymältään viestintäsovelluksia, jossa käyttäjän ja chatbotin dialogi esitetään viestiketjun tapaan, näyttämällä käyttäjän syötteet näytön laidassa ja chatbotin vastakaisella laidalla (Araujo, 2018).

Chatbottien viestintäsovellusmainen käyttöliittymä ei ole sattumaa. McTearin (2017) mukaan viestintäsovellusmainen käyttöliittymä on toimiva ratkaisu chatboteille, sillä se kannustaa käyttäjiä interaktioon ja vuoropuheluun järjestelmän kanssa. Toinen merkittävä hyöty viestintäsovellusmaisuudesta on chatbottien integroinnin helpottuminen: Chatbotteja pystyy julkaisemaan suoraan jo olemassa olevilla viestintäsovelluksilla. Slackin ja Facebook Messengerin kaltaiset viestintäsovellukset tarjoavat alustan, jonne chatbottien julkaiseminen ja integrointi on helppoa (McTear, 2017). Klopfensteinin, Delpriorin, Malatinin ja Bogliolin (2017) mukaan chatbottien yleistymiseen eri viestintäsovellusalus-

toilla on useita eri syitä. Viestintäsovellusalustat ovat hyvin suosittuja ja niiden käyttö on ohittanut jopa sosiaalisen median käytön, joten niiden välityksellä voi potentiaalisesti tavoittaa suuriakin käyttäjämääriä. Käyttäjät ovat myös tottuneet käyttämään eri viestintäsovelluslustoja, joten ne ovat toimintaympäristöltään tuttuja ja intuitiivisia käyttäjille. Tuttu ympäristö ja yhtenäinen alusta madaltaa kynnystä chatbottien kokeilemiseen ja omaksumiseen (Klopfenstein ym., 2017).

3 IHMISEN JA TIETOKONEEN VÄLINEN VUOROVAIKUTUS

Tässä luvussa käsitellään ihmisen ja tietokoneen välistä vuorovaikutusta muutamasta eri näkökulmasta. Aluksi esitellään lyhyesti ihmisen ja tietokoneen välinen vuorovaikutusta yleisemmin tutkimusalana. Tämän jälkeen käydään läpi ihmisen ja chatbotin vuorovaikutuksessa esiintyviä muutoksia ja haasteita aikaisempaan käsitykseen ihmisen ja tietokoneen välisestä vuorovaikutuksesta. Lopuksi tuodaan esille, miten ihmiset reagoivat sosiaalisesti tietokoneisiin yhdistämällä niihin totuttuja sosiaalisia käytänteitä ja inhimillisiä ominaisuuksia.

3.1 Ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimus

Ihmisen ja tietokoneen välinen vuorovaikutus on 1960-luvulla syntynyt tieteenhaara, joka tutkii tietokoneiden suunnittelua, arviointia ja toteutusta, sekä näiden ympärillä ilmeneviä ilmiöitä (Hewett ym., 1992). Tutkimusalan tarkoitus on ymmärtää ihmisen ja tietokoneen suhdetta, jotta tietoteknisten laitteiden ja ohjelmistojen käytöstä voitaisiin tehdä mahdollisimman tehokasta ja mielekästä. Ihmisen hienomotoriikalla, kognitiolla ja persoonallisuudella on vaikutusta tietotekniikan käyttöön, ja tämän vaikutuksen paremmalla ymmärryksellä saadaan kokonaisvaltaisempi kuva ihmisen toiminnasta, niin tietotekniikan kuin biologiankin näkökulmasta.

Ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimukseen tärkeimpinä käsitteinä liittyvät käytettävyys ja käyttäjäkokemus. Nielsenin (2012) mukaan käytettävyydellä tarkoitetaan ohjelmiston käytön helppoutta ja toimivuutta. Hän jakaa ohjelmiston käytettävyyden viiteen eri osatekijään: tehokkuuteen, opittavuuteen, mielekkyyteen, muistettavuuteen ja virnehallintaan (Nielsen, 2012). Käyttäjäkokemus on sen sijaan käytettävyyttä kokonaisvaltaisempi käsite, joka ohjelmiston käytön lisäksi pitää sisällään käytön aikana käyttäjälle syntyvät tunteet ja kokemukset (Hassenzahl & Tractinsky, 2006). Käytettävyys ja

käyttäjäkokemus ovat paljon tutkittuja ilmiöitä, joiden ymmärtämisellä on tärkeä osa tietotekniikan ja ohjelmistojen suunnittelussa ja kehityksessä.

Tietokoneita käytetään nykyään tavallisesti graafisten käyttöliittymien kanssa, jonka takia ihmisen ja tietokoneen vuorovaikutusta on tutkittu paljon niihin liittyen. Graafisessa käyttöliittymässä ihmisen ja tietokoneen vuorovaikutus perustuu ihmisen fyysisen syötteen aikaansaamaan reaktioon, joka näkyy muutoksena tietokoneen näytöllä (Dillon, 2003). Tämä vuorovaikutus on ihmiselle hyvin intuitiivista, sillä visuaalisuus mahdollistaa ihmiselle tuttujen käytänteiden liittämisen tietokoneen käyttöä auttamaan. Ikonit, osoittimet ja ikkunat auttavat hahmottamaan käytettävää järjestelmää ja esimerkiksi kansioiden ja roskakorin toiminnat muistuttavat toisiaan niin tietokoneella kuin fyysisessä maailmassa. Vuorovaikutus tietokoneen kanssa on säilynyt lähes muuttumattomana vuosikymmenien ajan, vaikka käytettävät teknologiat ja tietotekniset laitteet ovatkin kehittyneet sinä aikana merkittävästi (Beaudoin-Lafon, 2004). Graafisen käyttöliittymän pitkän historian takia, suuri osa ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimuksesta liittyy vahvasti sen ympärillä ilmeneviin ilmiöihin.

3.2 Vuorovaikutus chatbotin kanssa

Ihmisen ja chatbotin välinen vuorovaikutuksessa on selkeitä eroja totuttuun, graafisiin käyttöliittymiin perustuvaan vuorovaikutukseen. Chatbottien kanssa ihmisen ja tietokoneen välinen vuorovaikutus lähenee käyttäjän ja tietokoneen välistä interaktiivista dialogia, jolloin käyttäjän syöte ei entiseen tapaan aiheutakaan yksiselitteisesti jonkin toiminnon suorittamista näytöllä (McTear, 2017). Puhutun kielen käyttämisen ja interaktiivisuuden kautta, vuorovaikutus chatbotin kanssa voi vaikuttaa samankaltaiselta kuin vuorovaikutus oikeiden ihmisten kanssa (Gnewuch ym., 2017). Oviatt ja Cohen (2000) huomauttavat, että multimodaalinen vuorovaikutus mahdollistaa teknologioiden hyödyntämisen laajemmalle yleisölle. Multimodaalisella vuorovaikutuksella tarkoitetaan monien eri vuorovaikutuskanavien hyödyntämistä, joita voivat olla esimerkiksi tavanomainen hiiri ja näppäimistö, kosketusnäyttö, puhekomennot tai eleiden hyödyntäminen. Uudet tavat teknologioiden hyödyntämiseen parantavat niiden esteettömyyttä ja saavutettavuutta, helpottaen palvelujen käyttöä esimerkiksi näkövammaisille ja vähemmän tietotekniikkaa käyttäville (Oviatt & Cohen, 2000). Ihmisen ja chatbotin välinen vuorovaikutus sisältää siten uusia mahdollisuuksia tietotekniikan hyödyntämiselle eri tarkoituksissa.

Chatbottien dialogiin perustuva vuorovaikutus luo kuitenkin myös uusia haasteita ihmisen ja tietokoneen väliselle vuorovaikutukselle. Følstad ja Brandtzæg (2017) mukaan dialogin ollessa vuorovaikutuksen keskipisteessä, chatbottien kanssa käydyistä keskusteluista itsessään tulee oma suunnittelu-kohteensa chatbottien kehityksessä. Myös McTear (2018) korostaa keskustelujen suunnittelun merkitystä. Hän huomauttaa, että luonnollisen ja intuitiivisen vuorovaikutuksen luomiseksi, IT-alan osaajien tulisi omaksua ja hyödyntää kie-

litieteiden jo tuottamaa tietoa keskustelun rakenteesta ja mekaniikoista. Keskusteleavan vuorovaikutuksen suunnitteluun liittyy monia aiheen tutkimukselle uusia kysymyksiä, kuten miten käyttäjä sitoutetaan vuorovaikutukseen chatbotin kanssa, miten interaktion laatua voidaan mitata ja kuinka ihmismäinen chatbotin tulisi olla (McTear, 2018). Cameron ym. (2018a) korostavat myöskin chatbottien dialogin suunnittelun tärkeyttä. Keskustelujen suunnittelu ja rakentaminen ovat chatbottien kehittämisen haastavin osuus, mutta edellytys onnistuneen vuorovaikutuksen luomiselle (Cameron ym., 2018a). Følstad ja Brandtzæg (2017) argumentoivat keskusteleavan vuorovaikutuksen siirtävän painotuksen järjestelmän käytön selittämisen ja opettamisen sijaan käyttäjän tavoitteiden tunnistamiseen. Heidän mukaansa käyttöliittymäsuunnittelun ajatusmaailmasta tulee siirtyä palvelusuunnitteluun, jossa järjestelmän toiminta suunnitellaan enemmänkin käyttäjän ympärille eikä päinvastoin. (Følstad & Brandtzæg, 2017). Ihmisen ja chatbotin vuorovaikutuksen sosiaalinen elementti pakottaa katsomaan ihmisen ja tietokoneen välistä vuorovaikutusta uudesta näkökulmasta. Totutusta poikkeava vuorovaikutustapa kannustaa lähestymään teknologioita uusilla tavoilla, joka haasteidensa lisäksi luo kuitenkin pohjaa uusien innovaatioiden syntymiselle.

Sosiaalisen interaktion lisäksi haastetta ihmisen ja chatbotin vuorovaikutukselle lisää myös chatbottien tekninen toteutus. Vaikka chatbottien tekninen toteuttaminen on nykyisin suhteellisen helppoa, näiden toteutusten toimivuus ei ole mitenkään varmaa (Moore, Arar, Ren & Szymanski, 2017). McTear (2002) nostaa esille, että chatbottien kehittämisessä yhdistyvät monet suhteellisen uudet teknologiat, joiden käyttämisestä yhdessä ei ole hirveästi kokemusta. Uusien teknologioiden oppiminen vaatii aina aikaa ja kokemusta niiden käytöstä. Myös Cameron ym. (2018a) huomioivat chatbottien kehittämiseen olevan olemassa useita eri rajapintoja ja teknologioita, joista käyttötarkoitukseen sopivimman valitseminen ei ole aina selkeää. McTearin (2018) huomauttaakin yleisten standardien puutteen olevan yksi tämän hetken isoimmista ongelmista chatbottien kehittämisessä. Eri teknologioiden ja rajapintojen toiminnot ja nimeämiskäytännön vaihtelevat suuresti toistensa suhteen, joka hankaloittaa toimimista eri teknologioiden kanssa. Myös kehittäjien tietous, varsinkin luonnollisen kielen prosessoinnin suhteen, rajoittaa saatavilla olevien työkalujen tehokasta hyödyntämistä. Yhtenäisten standardien puutteen vuoksi chatbottien kehittämisessä nojaututaan tutkimuksen ja perusteltujen päätösten sijaan enemmänkin intuitioon (McTear, 2018). Teknologioiden osaamisen lisäksi, Cameron ym. (2018a) huomauttavat tarpeen myös totuttujen kehitysmenetelmien ja -prosessien muokkaamiseen chatbottien kehitystä varten. Heidän mukaansa kehitys ja testaus ovat yhdistyneet vahvemmin chatbotteja kehitettäessä, verrattuna tavanomaisempaan sovelluskehitykseen (Cameron ym., 2018a). Toimintatapojen muutoksen tarpeen huomioi myös Holmes ym. (2019), joiden mukaan tavanomaiset käytettävyyden ja käyttäjäkokemuksen mittarit eivät sovellu chatbottien arviointiin, jonka takia uusien arviointimetodien kehittämien on välttämätöntä chatbottien arviointia varten.

Chatbotin kehittämisen tekniset ja käyttökokemuksen haasteet heijastuvat myös ihmisen ja chatbotin väliseen vuorovaikutukseen. Suurimpia ongelmia ihmisen ja chatbotin vuorovaikutuksessa ovat järjestelmän toiminnallisuuden heikko välittyminen käyttäjälle (Luger & Sellen, 2016; Hill ym., 2015; Zamora, 2017; Chen & Wang, 2018), sekä järjestelmän kyvyttömyys vastata käyttäjien inhimillisiin ennako-odotuksiin (Luger & Sellen, 2016; Zamora, 2017; Portela & Granell-Canut, 2017; Jain ym., 2018). Chatbottien interaktiivisuus ja keskustelunomaisuus antaa käyttäjille vapauden käyttää järjestelmiä haluamallaan tavalla, mutta samalla se kuitenkin poistaa järjestelmien käyttäjiä ohjaavat raamit ja ohjenuorat. Tämä hämärtää chatbottien toiminnallisuuden rajoja, jonka takia ne saatetaankin nähdä yhtenä kokonaisuutena, eikä useita toimintoja sisältävänä ohjelmistona. Tätä ajatusta tukee Jainin ym. (2018) huomio, että chatbotin epäonnistuminen yhdessäkin toiminnassa vaikuttaa negatiivisesti koko chatbotista ja sen käytön jatkamisesta syntyvään mielikuvaan (Jain ym., 2018). Ihmisen ja chatbotin vuorovaikutuksen sosiaalinen, tekninen ja toiminnallinen vuorovaikutus eroaa siten suurestikin tietotekniikan aiemmasta käytöstä.

3.3 Tietokoneet ovat sosiaalisia toimijoita

Ihmisen ja chatbotin vuorovaikutuksen muistuttaessa osittain ihmisten välistä vuorovaikutusta, herääkin kysymys, miten ihmiset tulkitsevat sosiaalista vuorovaikutusta tietokoneiden kanssa? Tietokoneet ovat sosiaalisia toimijoita (engl. *Computers Are Social Actors, CASA*) on 90-luvulla kehittynyt paradigma, jonka mukaan ihmiset kohtelevat tietokoneita sosiaalisina olentoina (Nass & Moon, 2000). Varsinkin ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimuksessa, *CASA*-paradigmaa pidetään teoreettisena pohjana selittämään ihmisen sosiaalisia reaktioita tietokoneita kohtaan (Nass, Steuer & Tauber, 1994). Paradigman ajatuksena on, että tietokoneissa on jokin piirre tai ominaisuus, niin sanottu sosiaalinen vinkki, joka aikaansaa sosiaalisen reaktion käyttäjissä tietokonetta kohtaan (Nass & Moon, 2000; Fogg, 2002). Sosiaalisia vinkkejä voidaan käyttää apuna tietokoneiden sosiaalisia reaktioita synnyttävien ominaisuuksien hahmottamisessa, jolloin ihmisen ja tietokoneen välisestä vuorovaikutuksesta voidaan tehdä syvempiä havaintoja.

Ihmisille ovat tottuneet analysoimaan vuorovaikutustilanteissa vuorovaikutuskumppaniaan ja tekemään havainnoistaan tulkintoja vuorovaikutuskumppanin erilaisista sosiaalisista ominaisuuksista. Nass ja Moon (2000) selittävät tätä ilmiötä evoluutiolla ja miten aikaisemmin sosiaalista käytöstä esiintyi ainoastaan muissa ihmisissä. Tämän takia ihmiset hakevat ja tunnistavat tietokoneista sosiaalisia vinkkejä, jotka selittävät niiden sosiaalista käytöstä ja sosiaalista roolia (Nass & Moon, 2000). Sosiaalisten vinkkien tulkitseminen on ihmisille niin ominaista, että sitä tapahtuu automaattisesti ihmisen alitajunnassa ilman, että ihmiset välttämättä edes tiedostavat reagoivansa sosiaalisesti tietokoneeseen (Nass ym., 1994; Nass & Moon, 2000). Esimerkiksi ihmisten on todettu tiedostamattaan liittävän sukupuolistereotyyppioita tietokoneeseen siinä käyte-

tyn puheäänien perusteella, samalla kuitenkin kieltäen tehneensä niin (Nass, Moon & Green, 1997).

Sosiaalisia reaktioita synnyttäviä sosiaalisia vinkkejä on erilaisia ja ne ilmenevät eri asioista. Fogg (2002) tunnistaa viisi erityyppistä sosiaalisten vinkkien kategoriaa, jotka aikaansaavat ihmisen kokemaan tietokoneen sosiaalisena toimijana (taulukko 1). Hänen mukaansa esimerkiksi pelkkä tietokoneen kuvaaminen tiimitoveriksi saa ihmiset kokemaan sen älykkäämmäksi ja hyödyllisemmäksi, verrattuna identtiseen tietokoneeseen ilman assosioituja ominaisuuksia. Sosiaalisten vinkkien hyödyntäminen lisää ihmisten sosiaalista käyttäytymistä tietokoneiden kanssa, jonka vuoksi niitä voidaan hyödyntää tietoteknisten laitteiden suunnittelussa ja kehittämisessä (Fogg, 2002).

TAULUKKO 1 Tietokoneen sosiaalista läsnäoloa herättävät sosiaaliset vinkit (Fogg, 2002)

<i>Sosiaalinen vinkki</i>	<i>Esimerkki</i>
Fyysinen	Kasvot, keho, liikkeet
Psykologinen	Persoonallisuus, tunteet, empatia
Kieli	Puhuttu kieli, interaktiivinen kielenkäyttö
Sosiaalidynamiikka	Yhteistyö, vuorovastaisuus, kehuminen
Sosiaaliset roolit	Tiimitoveri, vastustaja, opettaja, ohjaaja

Fogg (2002) huomauttaa, että samalla kun sosiaalisten vinkkien hyödyntäminen voi onnistuessaan parantaa vuorovaikutusta tietokoneen kanssa, voi se yhtä hyvin epäonnistuessaan aiheuttaa negatiivisia reaktioita. Hän huomauttaa sosiaalisten vinkkien hyödyntämisen soveltuvan parhaiten viihteellisiin tai opetuskäyttöön tarkoitettuihin sovelluksiin. Sen sijaan tehtävän tehokkuutta parantaville sovelluksille hän ei näe sosiaalisten vinkkien hyödyntämisen olevan mielekästä. Tästä esimerkkinä hän mainitsee bensa-aseman tankkausautomaatin, jonka toiminnallisuuden hän kuvaa tarvitsevan yltää vain maksun suorittamiseen, eikä sosiaalisten piirteiden lisääminen toisi mitään lisäarvoa tähän tehtävään (Fogg, 2002).

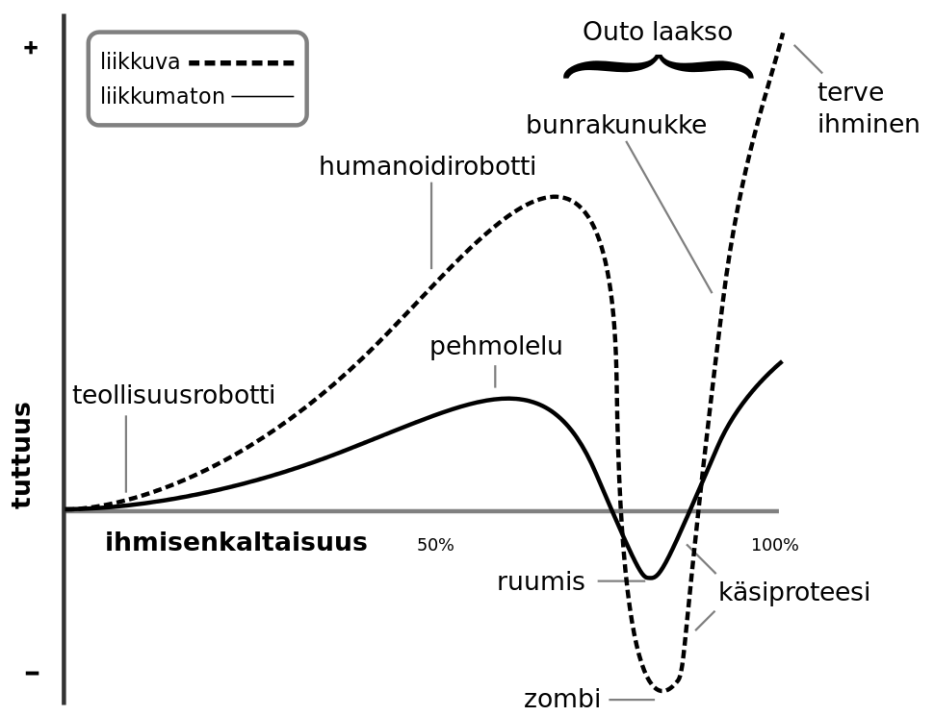
Tietokoneiden synnyttämien sosiaalisten reaktioiden yhteydessä esiintyy usein myös termi antropomorfismi. DiSalvo ja Gemperle (2003) määrittelevät antropomorfismin inhimillisten ominaisuuksien liittämiseksi ihmisiin lukeutumattomiin objekteihin. Antropomorfismia ovat esimerkiksi robotti-imurin nimeäminen tai tuttu piirrettyhahmo Mikki Hiiri, jossa kummassakin esimerkissä elottomalle asialle on annettu inhimillisiä ominaisuuksia: robotti-imurille nimi ja Mikki Hiirelle nimen lisäksi sosiaalinen ympäristö ja elämä. Antropomorfismin voidaan nähdä syntyvän ihmisen reaktioista aikaisemmin käsiteltyihin sosiaalisiin vinkkeihin.

Antropomorfismin syntyminen on osittain hämärän peitossa, mutta sen muodostumista on silti tutkittu ja spekuloitu. DiSalvon ja Gemperlen (2003) mukaan sosiaalitieteissä on useita eri teorioita antropomorfismin taustoihin.

Esimerkiksi tuttuusteorian mukaan ihmisellä on tarve selittää tuntemattomia asioita tuttujen asioiden kautta. (DiSalvo & Gemperle, 2003). Sen sijaan Epley, Waytz ja Cacioppo (2007) selittävät antropomorfismin kokemista kolmen eri tekijän kautta. Ensimmäinen tekijä on antroposentrisen tiedon saatavuus ja sen sovellettavuus, eli samakaltaisuuden tunteminen ja itsensä peilaaminen toisen toimintaan. Toinen tekijä on motivaatio selittää ja ymmärtää toimintaa, jolla tarkoitetaan ihmisen tarvetta selittää tuntematonta ja pitää tilanne hallinnassa. Kolmantena tekijänä on sosiaalinen motivaatio, eli halu sosiaaliseen kontaktiin ja yhteyteen (Epley ym., 2007). Vaikka antropomorfismin muodostuminen ei olekaan täysin selvää, sen voidaan todeta ilmiönä olevan todellinen.

Antropomorfismin syntymisen lisäksi, ei sen kokeminen myöskään ole yksiselitteistä, vaan siitäkin on risteäviä näkemyksiä. CASA-paradigman mukaan antropomorfismin kokeminen on ihmisen tietoinen päätös, sillä henkilön täytyy oikeasti uskoa asian tai esineen olevan elollinen kokeakseen antropomorfismia (Nass ja Moon, 2000). Kim ja Sundar (2012) esittävät kuitenkin vastakkaisen väitteen, sillä heidän mukaansa ihmiset tietoisesti kieltävät kokevansa antropomorfismia, mutta kuitenkin tiedostamattaan tekevät niin. Heidän tutkimuksensa mukaan hyvinkin yksinkertaiset antropomorfistiset vihjeet, kuten piirroshahmo nettisivulla, sai käyttäjissä aikaan edellä mainitun reaktion (Kim & Sundar, 2012). Tätä tukee myös aiemmin mainittu esimerkki sukupuolistereotyypioiden liittämistä tietokoneisiin, käyttäjien sitä tiedostamattaan (Nass, Moon & Green, 1997). Oli antropomorfismi tiedostamatonta tai ei, sen kokeminen on kuitenkin ihmisille ominaista ja yleistäkin. Antropomorfismilla voidaan siis todeta olevan vaikutusta tietokoneohjelmien käyttökokemukseen.

Antropomorfismin vaikutukset ihmisten kokemuksiin voivat olla sekä positiivisia että negatiivisia. DiSalvon ja Gemperlen (2003) mukaan, antropomorfismia ymmärtämällä ja sen suunnitelmallisella käytöllä voi olla positiivinen vaikutus tuotteeseen tai palveluun. Esimerkiksi vartalonmuotoinen hajuvesipullo välittää mielikuvaa aistillisuudesta ja seksuaalisuudesta, saaden tuotteen tuntumaan puoleensavetävämmältä. Ilmiötä voidaan hyödyntää myös tietotekniikassa kannustamaan käyttäjää haluttuun toimintaan, siten parantaen tuotteen tai palvelun käyttökokemusta (DiSalvo & Gemperle, 2003). Antropomorfismin negatiivisia vaikutuksia selittää Morin (1970) Oudon laakson -teoria (engl. uncanny valley). Morin mukaan esineen tai asian ihmismäisyydessä löytyy piste, jossa esine tai asia ei tunnu enää tutulta ja sen takia aiheuttaa epä-tietoisuuden tunnetta ihmiselle (ks. kuvio 2). Mahdollinen selitys Oudon laakson -teorialle liittyy ihmisen suojeleuvaistoon, sillä esimerkiksi kankealiikkeinen robotti voi alitajuntaisesti muistuttaa enemmän kuollutta ruumista kuin ihmistä, jolloin siitä erossa pysyminen on luontainen reaktio (Mori, 1970). Antropomorfismin negatiiviset vaikutukset vaikuttavat kuitenkin sijoittuvan suhteellisen kapealle skaalalle antropomorfismia, jonka takia positiivisten vaikutusten tavoittelu suunnittelulla voi olla ihan varteen otettava vaihtoehto palveluiden kehittämisessä.



KUVIO 2 Morin outo laakso (Pafcu, 2007)

4 CHATBOTTIEN SYNNYTTÄMÄT SOSIAALISET REAKTIOT KÄYTTÄJISSÄ

Aikaisemmassa luvussa käsiteltiin ihmisen ja tietokoneen välisen vuorovaikutuksen sosiaalisia piirteitä ja niiden muodostumista. Tässä luvussa aihetta tarkastellaan kohdistetummin ihmisen ja chatbotin vuorovaikutuksen näkökulmasta. Aluksi esitellään mitä odotuksia käyttäjillä on chatbottien heissä synnyttämille sosiaalisille reaktioille ja sen ihmismäiselle toiminnalle. Tämän jälkeen käsitellään chatbottien eri piirteitä ja ominaisuuksia, joiden välityksellä chatbotteihin kohdistuvat sosiaaliset reaktiot välittyvät käyttäjälle. Lopuksi esitetään huomioita chatbottien synnyttämien sosiaalisten reaktioiden tulkintaan liittyen.

4.1 Ihmisen ja chatbotin väliseltä vuorovaikutukselta odotetut sosiaaliset reaktiot

Kuten mihin tahansa muuhun asiaan, myös chatbotteihin kohdistuu erilaisia ennakko-odotuksia. Chatbottien ympärillä pyörivässä keskustelussa chatbotteihin liitetään monia eri termejä, joista yksi yleisimmistä on varmastikin ”tekoäly”. Kyseinen termi jo itsessään saa chatbotin kuulostamaan älykkäältä, tietoisuuden omaavalta entiteetiltä, riippumatta siitä mikä onkaan chatbotin todellinen toiminnallisuus ja tarkoitus. Bond ym. (2019) tuovat ilmi tekoälyn ja tietokoneen ihmismäisten piirteiden herättävän ihmisissä epävarmuutta ja jopa pelkoa. Esimerkiksi työpaikan menetys on yksi yleisistä tekoälyyn liittyvistä huolistista, joka osaltaan lisää uusiin teknologioihin kohdistuvaa varauksellisuutta (Bond ym., 2019). Epävarmuus ja pelko eivät kuitenkaan ole ainoita ennakkoluulojen herättämiä tunteita chatbotteja kohtaan. Jain ym. (2018) huomioivat joidenkin ihmisten kohdistavan uusiin teknologioihin myös toiveikkaita ja myönteisiä odotuksia. Jotkut ihmiset itseasiassa odottavat chatbottien olevan ihmismäisiä ja haluavat vuorovaikuttaa chatbottien kanssa samoin, kuin ihmisten välisen vuorovaikutuksessa on tapana (Jain ym., 2018). Chatbottien vuorovaikutuksellinen käyttötapa varmasti vahvistaa odotuksia chatbottien inhimillisistä ominaisuuksista, mutta se ei tietenkään takaa chatbotin vastaavan näihin odotuksiin.

Yleinen ihmisten asettama odotus chatboteista syntyviin sosiaalisiin reaktioihin on, että chatbottien odotetaan omaavan sosiaalista älykkyyttä ja kykyä toimia sosiaalisissa tilanteissa. Tämän huomasi esimerkiksi Zamora (2017) tutkimuksessaan, johon osallistuneet käyttäjät toivoivat chatbotin kykenevän tarjoamaan henkistä tukea ja toimimaan keskustelukumppanina. Chatbotista haetaan objektiivista hahmoa, jolle voi keskustella mielenkiinnon kohteista ja henkilökohtaisemmistakin aiheista, ilman ihmisten väliseen vuorovaikutukseen liittyviä sosiaalisia paineita (Zamora, 2017; Thies, Menon, Magapu, Subramony & O'Neill, 2017). Ho, Hancock ja Miner (2018) arvelevat sosiaalisen interaktion hakemisen chatbotilta juurtavan ihmisen luontaisesta tarpeesta hyväksyntään ja yhteenkuuluvuuden tunteeseen. Ajatus chatbotista emotionaalisenä keskustelukumppanina ei olekaan niin kaukaa haettu, sillä Ho ym. tunnistivat vuorovaikutuksen chatbotin kanssa luovan samat emotionaaliset, vuorovaikutukselliset ja psykologiset hyödyt kuin vuorovaikutuksen ihmisen kanssa (Ho ym., 2018). Ehkä edellä mainittu onkin osatekijänä sille, että chatbotteja onkin sovellettu ja kehitetty onnistuneesti esimerkiksi mielenterveyden hoitoon liittyvissä tarkoituksissa ja ympäristöissä (Cameron ym., 2018b). Sosiaalisten suhteiden luominen ja ylläpitäminen tapahtuu vuorovaikutuksen kautta, jonka vuoksi näiden asioiden hakeminen myös chatbottien vuorovaikutuksesta on ymmärrettävää.

Sosiaalisen älykkyyden lisäksi, myös chatbottien persoonallisuuteen ja luonteeseen kohdistuu ennako-odotuksia. Ghoshin ja Pherwanin (2015) tutkimukseen osallistuja odottivat virtuaaliavustaja-chatbotin vastaavan oikeaa avustajaa niin toiminnaltaan kuin käytökseltään. Tutkimuksessa huomattuja chatbottien toivottuja ominaisuuksia olivat proaktiivisuus, tehtäväkeskeisyys sekä kekseliäisyys, kuitenkin ilman liiallista sympaattisuutta. Tehtäviä tehdessä chatbottien haluttiin olevan hillitty, avulias ja toimivan käyttäjän alaisena. Sen sijaan, tehtävien ulkopuolella chatbotin oli hyväksyttävää tuoda ilmi myös omia mielipiteitään ja vapaampia reaktioita (Ghosh & Pherwani, 2015). Chatbotin havaittu olemus ja persoonallisuus vaikuttaa myös sen käyttäjien suuntaan, sillä persoonallisuudeltaan vakavaa ja itsevarmaa chatbottia kuunnellaan herkemmin ja sille uskoudutaan luottavamemmin, verrattuna olemukseltaan rennompaan chatbottiin (Li, Zhou, Yang & Mark, 2017). Chatbottien toimintaan voidaan tulkita kohdistuvan hyvinkin samoja odotuksia, kuin ihmisiinkin. Yleisesti maltillisten ja asiallisten piirteiden voidaan todeta olevan chatbottien odotetuimpia persoonallisuuspiirteitä. Chatbottien odotettu toiminta vaikuttaa seuraavan hyvinkin vahvasti yleisiä sosiaalisia normeja.

4.2 Chatboteissa ilmenevät sosiaaliset vinkit

Millä tavalla chatbotin ihmismäisyys sitten muodostuu? Aiemmin käsiteltiin CASA-paradigmaa ja tietokoneiden sisältämiä sosiaalisia vinkkejä, joihin ihmiset reagoivat sosiaalisesti. Tähän pohjautuen, Feine ym. (2019) tutkivat sosiaalisten vinkkien ilmenemistä chatboteissa. He tunnistivat yhteensä 48 eri sosiaa-

lista vinkkiä, jotka jakautuvat neljään eri kategoriaan. Tunnistetut sosiaaliset vinkit ja niiden kategoriat esitetään taulukossa 2. Näitä sosiaalisia vinkkejä voidaan hyödyntää chatbottien arvioinnissa, sekä ne mahdollistavat chatbottien ihmismäisyyden muodostumisen tarkastelua järjestelmällisemmin kuin ennen (Feine ym., 2019). Chatboteissa ilmenevien sosiaalisten vinkkien tunnistaminen on tulos osa ihmisen ja chatbotin välisen vuorovaikutuksen tutkimukselle. Chatbotteja arvioidaan itseasiassa enemmän niistä tulkittujen sosiaalisten vinkkien perusteella, kuin itse toiminnan tai käytön perusteella (Koh & Sundar, 2010). Tätä huomiota vahvistaa Tanin ja Liewin (2020) tutkimus, jonka mukaan käyttökontekstiin ja -ympäristöön sopivia sosiaalisia vinkkejä sisältävä chatbot koetaan uskottavammaksi, läheisemmäksi ja luotettavammaksi, kuin saman toiminnallisuuden omaava chatbot ilman kohdistettuja sosiaalisia vinkkejä. Feinen ym., (2019) tunnistamia sosiaalisten vinkkien kategorioita tarkastellaan ja analysoidaan tarkemmin seuraavissa aliluvuissa.

TAULUKKO 2 Chatbotin sosiaalisten vinkkien luokittelu (Feine ym., 2019)

<i>Kategoria</i>	<i>Alakategoria</i>	<i>Sosiaalinen vinkki</i>	
Verbaalinen	Sisältö	Keskustelun ohjaus Selitys Tervehdys Vitsi Myötämielisyys Ylistys	Menneeseen viittaaminen Itsestä kertominen Kysymykset itsestä Small talk Kiittäminen Vinkit ja neuvot
	Tyyli	Lyhenteet Formaalius Sanastollinen laajuus	Virkkeiden monimutkaisuus Kielen painokkuus
Visuaalinen	Kinesiikka	Käsimerkit ja -eleet Silmänliikkeet	Ilmeet Pään liikkeet
	Proksemiikka	Taustakuva	Vuorovaikutuksellinen etäisyys
	Agentin ulkoasu	2D-/3D-visualisointi Ikä Viehättävyys Vaatetus Väritys	Realistisuus Kasvonpiirteet Sukupuoli Nimi Fotorealismi
	Tietokonevälitteinen viestintä	Hymiöt	Kirjasintyyppi
Auditiivinen	Äänen ominaisuudet	Sukupuoli Sävelkorkeus	Tempo Äänenvoimakkuus
	Ääntely	Nauru Murahdus ja voihe	Haukottelu Vokaalinen segregaatio
Näkymätön	Kronemiikka	Aloitteen tekeminen	Vasteaika
	Haptiikka	Taktiilinen kosketus	Lämpötila

4.2.1 Verbaaliset vinkit

Feinen ym., (2019) ensimmäinen chatboteista tunnistama sosiaalisten vinkkien kategoria on verbaaliset vinkit. Verbaalisilla vinkeillä viitataan sosiaalisiin vinkkeihin, jotka välittyvät tai muodostuvat sanojen ja viestien sisällön perusteella. Feine ym. seuraavat Waltherin (2008) ajatusta siitä, että kielen synnyttämät sosiaaliset reaktiot riippuvat viestin tyylistä ja sanallisesta sisällöstä: Saman asian voi ilmaista monella eri tapaa ja tällä esitystavalla on vaikutusta siihen, millaisena viesti vastaanotetaan. Tämän takia verbaaliset jakautuvat vielä tarkemmin sisällön ja tyylin alakategorioihin. (Feine ym., 2019).

Viestien sisällöllä ja ilmaisutyyllillä on todettu olevan vaikutusta chatbotteihin yhdistettävien ominaisuuksien syntymiseen. Dryerin (1999) on huomioi-

nut ihmisten liittävän persoonallisuuksia tietokoneisiin samaan tapaan kuin ihmisiin: tulkitsemalla sen kielenkäyttöä, sanamuotoja ja muita sosiaalisia vinkkejä. Chatbotin havaittu persoonallisuus ja älykkyys saa sen tuntumaan lähemmältä käyttäjälle (Lee, Lee & Sah, 2019). Chaves ja Gerosa (2020) argumentoivat, että se miten chatbot reagoi käyttäjän syötteisiin eri tilanteissa, vaikuttaa chatbotin koettuun sosiaaliseen ja keskustelulliseen älykkyyteen. Hei huomioivat tutkimuksessaan, että vaikka chatbotin vuorovaikutukselta odotetaan ihmisen tasoon verrattavaa sosiaalista älykkyyttä, voi chatbot vaikuttaa älykkäältä vaikkei se kuitenkaan täydellisesti ymmärtäisi kaikkea sille sanottua. Esimerkiksi virhetilanteissa chatbotin ei ole tarpeellista ymmärtää täysin mitä tapahtuu, vaan pelkkä tilanteen ohjaaminen takaisin haluttuun toimintaan riittää vahvistamaan mielikuvaa sen älykkyydestä (Chaves & Gerosa, 2020). Tällaisen kontekstiin sopivan toiminnan merkitystä tukee myös Sundarin, Bellurin, Ohin, Jian, ja Kimin (2016) tutkimus, jossa kontekstin kannalta yhtenäisten ja siihen sopivien vastausten todettiin vahvistavan chatbotista tulkittuja ihmismäisiä ominaisuuksia. Ihmismäisen mielikuvan syntyä voi vahvistaa myös osaltaan pienet virheet, kuten satunnaiset kirjoitusvirheet, jotka saavat chatbotin vaikuttamaan samaistuttavammalta (Dryer, 1999; Warwick & Shah, 2016). Verbaalisten vinkkien luoma kuva chatbottien antropomorfistisista ominaisuuksista vaikuttaa olevan vahvasti kontekstisidonnaista. Käyttäjät arvioivat chatbotin viestien sisältöä ja tyyliä tietyn kontekstin näkökulmasta, jonka rajoissa toimiminen riittää mielikuvan älykkyydestä muodostumiseen.

Verbaalisten vinkkien muodostamaan kuvaan chatbotin ihmismäisyydestä vaikuttaa vahvasti myös chatbotin tuottaman dialogin interaktiivisuus. Go ja Sundar (2019) huomioivat chatbotin koetun ihmismäisyyden kasvavan sitä enemmän, mitä interaktiivisempaa sen kanssa käyty dialogi on. He argumentoivat interaktiivisen dialogin lisäävän chatbotin koettua sosiaalista läsnäoloa, samaistuttavuutta ja keskustelun dynaamisuutta (Go & Sundar, 2019). Käyttäjien kiintymystä ja tunnesidettä chatbottia kohtaan lisää Portelan ja Granell-Canutin (2017) mukaan varsinkin se, jos chatbot muistaa jonkin käyttäjän aikaisemmin sanoman asian. Warwick ja Shah (2016) kuitenkin huomauttavat, ettei interaktiivisen ja keskusteluun mukautuvan dialogin ylläpitäminen ole helppoa, eivätkä nykyiset chatbotit onnistu ylläpitämään ihmismäisyyden illuusioita pidempään jatkuvissa keskusteluissa. Ihmisen ja chatbotin vuorovaikutuksen interaktiivisuus on looginen edellytys chatbotin ihmismäiseltä vaikuttavan vuorovaikutuksen synnylle. Vastavuoroinen dialogi kuitenkin jatkuessaan lisää siitä tulkittavia sosiaalisia vinkkejä, joka on varmaankin syynä chatbottien epäonnistumiseen ihmismäisen vuorovaikutuksen ylläpitämisessä. Ennemmin tai myöhemmin käyttäjät huomaavat chatbottien toiminnallisuuksien rajallisuudet ja mielikuva chatbottien sosiaalisista ominaisuuksista ja kyvykkyyksistä haalenee.

Verbaalisia vinkkejä arvioidaan myös niistä tulkitun huumorin ja sen käytön kannalta. Huumori ja vitsit ovat yleisiä chatbottien vuorovaikutuksessa käytettyjä sosiaalisia vinkkejä (Feine ym., 2019). Tämä ei ole yllättävää, sillä Jainin ym. (2018) mukaan ihmisen ja chatbotin vuorovaikutuksen leikinomaisuus

vaikuttaa usein positiivisesti siitä syntyvään käyttökokemukseen. Luger ja Sellen (2016) nostivat myös tutkimuksessaan esille, miten huumori ja viihdekäyttö ovat monelle käyttäjälle ensikosketus chatbottien käyttöön, jonka avulla chatbotin käytön oppiminen tapahtuu luontevasti hovin rinnalla. Humoristinen vuorovaikutus voi heidän mukaansa kuitenkin vääristää ja jopa yliampua chatbotin synnyttämää mielikuvaa sen ihmismäisten ominaisuuksien tasosta, asettaen odotuksen vuorovaikutuksen tasosta liian korkealle. Syynä tähän on se, että usein viihteelliset toiminnot ovat varta vasten kehitettyjä, eikä chatbot tavallisessa käytössä ole yhtä eloisa ja interaktiivista (Luger & Sellen, 2016). Humoristinen vuorovaikutus voi olla hyvä chatbotteihin tutustumiseen, mutta sen haasteena on jo aikaisemmin esitellyt odotukset chatbotin ammattimaisesta käytöksestä sekä keskustelun interaktiivisuuden ylläpitäminen.

4.2.2 Visuaaliset vinkit

Feinen ym. (2019) visuaalisten vinkkien kategoriaan sisältyvät nonverbaaliset sosiaaliset vinkit, jotka ovat ihmisen visuaalisesti tulkittavissa. Näihin kuuluvat kinesiikka, proksemiikka, agentin ulkoasu ja tietokonevälitteinen viestintä. Kinesiikalla tarkoitetaan liikkeiden, eleiden ja ilmeiden kaltaisten visuaalisten muotojen synnyttämiä sosiaalisia vinkkejä. Proksemiikka koskee tilankäytöstä tulkittavia viestejä, jotka voivat ilmetä esimerkiksi havaitun sosiaalisen etäisyyden tai henkilökohtaisen tilan käytön perusteella. Agentin ulkoasu sisältää chatbotin visualisoinnista, kuten vaatetuksesta ja hahmon realismisuudesta syntyvät sosiaaliset vinkit. Viimeinen kohta, eli tietokonevälitteinen viestintä liittyy chatbotin käyttämään kirjasimeen ja mahdollisien emojiiden käyttöön. Vaikka viestien sisältö ja tyyli kuuluvat aiemmin käsiteltyihin verbaalisiin vinkkeihin, viestien esittämiseen käytetyt keinot ovat silti visuaalisesti tulkittavissa (Feine ym., 2019).

Visuaalisuus on ihmiselle luonnollinen tapa havainnoida ympäristöä, ja ihmiset hyödyntävät sitä paljon myös sosiaalisten tilanteiden tulkitsemisessa. Kuten Feine ym. (2019) huomioivatkin, ihminen on oppinut tunnistamaan monia sosiaalisia vinkkejä vuorovaikutuskumppaneidensa pelkästä elehdinnästä ja ulkonäöstä. Ihmisille muodostuu kuva chatbotin persoonallisuudesta jopa pelkästään sen elehdintää seuraamalla ja tulkitsemalla (Liu, Tolins, Fox Tree, Neff & Walker, 2015). Rehmin ja André (2005) mukaan ihmiset kohdistavat vuorovaikutustilanteessa samoja sosiaalisia normeja niin ihmisiin kuin tietokonehahmoihin. Joskin he huomauttavat tietokonehahmojen kanssa näiden sosiaalisten normien noudattaminen ei ole yhtä vahvaa kuin ihmisten seurassa. Tämä voi ilmetä esimerkiksi chatbotin tölläämisinä, jota pidettäisiin epäkohteliaana, mikäli tölläyksen kohteena olisi toinen ihminen (Rehm & André, 2005). Visuaalisten vinkkien voidaan nähdä olevan merkittävä osatekijä chatbottien synnyttämien sosiaalisten reaktioiden havaitsemisessa, joskaan niiden tulkinta ei ole yhtä vahvaa verrattuna toisesta ihmisestä tehtyihin tulkintoihin.

Visuaalisilla vinkeillä nähdään olevan vahva yhteys antropomorfismin kokemiseen chatbottia kohtaan. Go ja Sundar (2019) huomioivat antropomor-

fismilla olevan vaikutusta chatbotteihin liitettyihin tulkintoihin ja odotuksiin. Esimerkiksi yksinkertainen chatbotin nimeäminen on tehokas tapa saada chatbot näyttämään ihmismäiseltä (Go & Sundar, 2019; Feine ym., 2019; Araujo, 2018). Go ja Sundar (2019) kertovat antropomorfismin kokemisella olevan yleensä chatbottien käyttökokemusta parantava vaikutus, sen lisäämällä chatbotin ihmismäisyyden tuntua. He mainitsevat antropomorfismin avulla olevan mahdollista kompensoida pieniä vajeavaisuuksia ihmisen ja chatbotin välisessä vuorovaikutuksessa, jotka muuten voisivat nousta esiin herkemmin. He kuitenkin myös varoittavat, että chatbotin koetun ihmismäisyys kasvaessa, voivat käyttäjien siihen kohdistamat odotukset myös nousta. Korkeampiin odotuksiin vastaaminen on aina hankalampaa ja jos siinä epäonnistutaan, on sillä negatiivisia vaikutuksia koko chatbotista syntyvään mielikuvaan. (Go & Sundar, 2019). Visuaalisten vinkkien todetaan auttavan chatboteista nousevien antropomorfististen kokemusten syntyä, jopa hyvinkin pienien tai yksinkertaisten ärsykkeiden välityksellä. Tämä tarkoittaa sitä, että käyttäjälle chatbotista syntyviä sosiaalisia reaktioita voitaisiin mahdollisesti ohjata pienin, mutta harkituin vihjauksin.

Chatbotin sisältämien visuaalisten vinkkien realistisuudella on todettu olevan suuri vaikutus niistä syntyviin sosiaalisiin reaktioihin. Gongin (2008) argumentoi chatbotin synnyttämän antropomorfismin kokemisen olevan sitä vahvempaa, mitä realistisempia elementtejä chatbotin visualisoinnissa käytetään. Hänen mukaansa chatbotin visuaalisen realistisuuden kasvaessa, myös sen välittämien sosiaalisten vinkkien määrä kasvaa (Gong, 2008). Rincón-Nigro ja Deng (2013) huomioivat chatbotin visualisoinnin realistisuudella olevan positiivisia vaikutuksia chatbotin käyttökokemukseen. He havaitsivat esimerkiksi luonnollisen elehdinnän helpottavan käyttäjiä chatbotin omaksumisessa sekä itse käytössä (Rincón-Nigro & Deng, 2013). Myös Beun ym. (2003) tunnistivat realistisen ulkoasun vaikuttavan positiivisesti chatbottien käyttökokemukseen. He huomasivat tutkimuksessaan chatbotista koetun antropomorfismin lisäävän vuorovaikutuksen muistettavuutta, joka puolestaan heijastui käyttökokemukseen sitä parantamalla (Beun ym., 2003). Sopivan realistisuuden raja on kuitenkin häilyvä, sillä liian realistinen visualisointi voi aiheuttaa Morin (1970) Oudon laakson -teorian mukaisia negatiivisia tuntemuksia. Tätä tukee Ciechanowskin ym. (2018) tutkimus, jossa he tunnistivat visuaalisen avatarin aiheuttavan käyttäjissä vahvempia reaktioita, kuin pelkän tekstin käyttäminen chatbotissa. Heidän tutkimustulosten perusteella pelkkää tekstiä ja dialogia käyttävä chatbot tulkitaan usein positiivisemmin, kuin realistisen esitystavan sisällyttäminen (Ciechanowski ym., 2018).

Myös chatbotin käyttämän kirjasimen on havaittu vaikuttavan chatbotista syntyvien sosiaalisten reaktioiden syntymiseen. Candello, Pinhanez ja Figueredo (2017) vertailivat tutkimuksessaan konemaisen "ORC-A"-kirjasimen ja ihmisen käsialaa imitoivan "Bradley"-kirjasimen vaikutuksia chatbotin tulkitsemiseen. Molemmilla kirjasimilla chatbot tulkittiin konemaiseksi, joskin ihmisen käsialaa imitoiva Bradley-kirjasin koettiin negatiivisemmin. Ihmisen käsialaa muistuttava fontti koettiin epäilyttäväksi, sillä se tulkittiin tietokoneen

yritykseksi näyttäytyä ihmismäisempänä kuin oikeasti on. Kirjasimen vaikutus chatbotista syntyvään mielikuvaan kohdistuu kuitenkin vain ensivaikutelmaan ja heikkenee järjestelmän käytön myötä. (Candello ym., 2017). Käisialaa imitoivan tekstin herättämät epäilykset chatbottia kohtaan ovat mielenkiintoinen ilmiö. Siinä missä ihmisen käisialaa imitoiva kirjasin voi herättää negatiivisia tunteita, samalla avatarin kanssa ihmistä lähenevällä visuaalisuudella nähdään olevan yleisesti positiivisia vaikutuksia.

4.2.3 Audiitiiviset vinkit

Audiitiivisiin tekijöihin liitetään Feine ym. (2019) mukaan kaikki nonverbaaliset, äänistä syntyvät sosiaaliset vinkit. Audiitiivisiin vinkkeihin kuuluvat äänen ominaisuutta välittävät sekä ääntelyiden aiheuttamat vinkit. Äänen ominaisuudet sisältävät pysyviä ominaisuuksia, kuten äänen sävelkorkeuden skaala, sekä muuttuvia ominaisuuksia, kuten tempo ja äänenvoimakkuus. Erilaiset ääntelyt, kuten itku, nauru ja murahdukset välittävät myös tulkittavia sosiaalisia vinkkejä, vaikeivat itsessään sisällä mitään kielellistä sisältöä (Feine ym., 2019)

Audiitiivisiä vinkkejä tulkitaan niiden kontekstin perusteella. Lee ja Nass (2003) korostavat kuinka audiitiivisten vinkkien perusteella chatbotista syntyneen mielikuvan tulisi sopia sen käyttökohteeseen. Ihmiset tekevät äänen ja ääntelyiden perusteella monia oletuksia puhujasta ja voivat esimerkiksi luoda oletuksen puhujan persoonallisuudesta. Tietokoneen käyttämän äänen ollessa yhtenevä sen välittämän verbaalin viestin kanssa, koetaan se miellyttäväksi ja toimivaksi. Tietokoneen yhtenäinen ja käyttökontekstiin sopiva ääni lisää käyttäjän kokemaa sosiaalista läsnäoloa tietokonetta kohtaan (Lee & Nass, 2003). Audiitiiviset vinkit yhdistetään vahvasti sosiaaliin normeihin, jonka perusteella niiden sopivuutta kontekstiin arvioidaan. Kuten monessa muussakin asiassa, tutuus koetaan miellyttävänä.

4.2.4 Näkymättömät vinkit

Feine ym. (2019) tunnistivat chatboteissa myös sosiaalisia vinkkejä, joita ei voi nähdä tai kuulla. Näitä näkymättömiä vinkkejä ovat kronemiikkaan ja haptiikkaan liittyvät sosiaaliset vinkit. Kronemiikalla tarkoitetaan ajan ja ajankäytön roolia kommunikaatiossa, jota ilmenee chatboteissa esimerkiksi vasteajan ja viestien temmon kautta. Haptiikka liittyy sen sijaan kosketuksen tunteeseen, joka voi välittyä esimerkiksi mobiililaitteen värinän kautta (Feine ym., 2019).

Näkymättömät vinkit voivat vaikuttaa chatbotin synnyttämiin sosiaaliin reaktioihin hyvinkin huomaamattomin elein. Portela ja Granell-Canut (2017) huomioivat miten pieni viive chatbotin antamassa vastauksessa voi näyttäytyä käyttäjälle chatbotin älykkyytenä. Viive luo mielikuvan siitä, että chatbot miettii vastaustaan ja siten saa käyttäjässä aikaan sosiaalisen reaktion chatbotin älykkyydestä (Portela & Granell-Canut, 2017). Myös Gnewuchin, Moranin, Adamin ja Maedchen (2018) ovat sitä mieltä, että chatbotin vaihteleva vasteaika vaikuttaa positiivisesti chatbotista syntyvään mielikuvaan. He esittävät vasteajan dy-

naamisen muuttamisen viestien monimutkaisuuden perusteella lisäävän chatbotin synnyttämistä sosiaalisista reaktioista koettua ihmismäisyyttä ja sosiaalista läsnäolon tuntua. Vasteajan vaikutuksia ei kuitenkaan täysin vielä ymmärretä ja käyttäjien yksilölliset erotkin vaikuttavat niiden tulkinnassa (Gnewuch ym., 2018). Myös Portelan ja Granell-Canutin (2017) huomiot tukevat edellä mainittua. Toiset käyttäjät voivat tulkita vastauksen viiveen juuri älykkyyden merkkinä, kun taas toiset tulkitsevat sen järjestelmän tehottomuutena (Portela & Granell-Canut, 2017). Kuitenkin myös liian nopea vastaaminen voidaan kokea häiritsevänä (Ciechanowski ym., 2018). Näkymättömien vinkkien näkymättömyys onkin varmaan osittain syynä niiden tulkinnan subjektiivisuudelle, sillä niiden tunnistaminen on hankalampaa kuin esimerkiksi visuaalisten tai verbaalisten ominaisuuksien tunnistaminen.

4.3 Sosiaalisten vinkkien tulkintaan vaikuttavia tekijöitä

Chatbottien sisältämien sosiaalisten vinkkien synnyttämien sosiaalisten reaktioiden tulkintoihin vaikuttaa myös chatbottiin liittymättömiä tekijöitä. Vaikka Feine ym., (2019) tunnistivatkin sosiaalisten vinkkien olevan chatboteista syntyvien sosiaalisten reaktioiden taustalla, sosiaalisten vinkkien tulkinta voi vaihdella joissain määrin käyttäjien yksilöllisten ominaisuuksien mukaan. Käyttäjien sukupuolen, iän, kokemusten sekä halun käyttää teknologiaa on yleisesti todettu muokkaavan sitä, miten käyttäjät näkevät teknologian ja onko sen käyttö vaivan arvoista (Venkatesh, Thong & Xu, 2012). Yhtä lailla, käyttäjien teknologiaan kohdistamat odotukset ja niiden toteutuminen määrittävät omaksuuko käyttäjä kyseisen teknologian käytön vai ei (Venkatesh & Bala, 2008). Chatboteissa ilmenevien sosiaalisten vinkkien tulkinnanvaraisuus onkin varmasti suuri osatekijä ihmisen ja chatbotin välisen vuorovaikutuksen ongelmissa.

Käyttäjien subjektiiviset ominaisuudet voivat vaikuttaa chatbottien sosiaalisten vinkkien tulkintoihin, jonka takia sosiaalisten vinkkien vaikutuksien arviointi on hankalampaa ja arvaamatonta. Devaraj, Easley ja Crant (2008) nostavat esille, kuinka käyttäjien persoonallisuudet jo itsessään vaikuttavan järjestelmään kohdistuviin odotuksiin. Tämän lisäksi odotuksien chatbottien ihmismäisyydestä ja sen persoonallisuudesta on havaittu olevan kulttuurisidonnaisia, vaihdellen paikallisten tapojen, arvojen ja normien mukaan (Danielescu & Christian, 2018). Monissa tutkimuksissa on tullut ilmi, miten samankaltaisuuden katsotaan vetävän ihmisiä puoleensa, jonka vuoksi käyttäjät useimmiten suosivat heidän omien ominaisuuksien mukaisia piirteitä heijastavia chatbotteja (Dryer, 1999; Lee, Peng, Jin & Yan, 2006; Payne, Szymkowiak, Robertson & Johnson, 2013). Chatbotin samankaltaisuuden tuntemisella voidaan olettaa olevan positiivisia vaikutuksia sen samaistuttavuuteen ja yhteenkuuluvuuden tunteeseen. Samat sosiaaliset vinkit voidaan kuitenkin tulkita täysin eri tavalla riippuen tulkitsijan demografisista tekijöistä, jonka takia sosiaalisten vinkkien vaikutuksien tulkinnassa totuuksien löytäminen on hankalaa.

ymmärtää tai osaa toimia niiden mukaan (Hill ym., 2015; Luger & Sellen, 2016; Jain ym., 2018). Yksinkertainen kieli rajaa viestien mahdollista kompleksisuutta, jonka takia siitä välittyy myös vähemmän verbaalisia sosiaalisia vinkkejä. Zamora (2017) painottaa myös, että chatbotit toimivat kiitettävästi ainoastaan englanniksi, joten sen toiminta muilla kielillä voi olla paljonkin heikompaa. Luonnollisen kielen prosessointi on varmasti yksi suurimpia rajoitteita ihmisen kaltaisen vuorovaikutuksen muodostamisessa chatboteille. Ihmisten käyttämä kieli sisältää monimutkaisia viestejä, joiden replikointi koneellisesti ei ole helppoa. Chatbottien älykkyudesta voi muodostua hyvinkin tyhmä mielikuva, mikäli sen viestien taso ei vastaa sujuvaa kieltä.

Merkittävä ja ehkä yllättäväkin tekijä chatbotin synnyttämiin sosiaalisiin reaktioihin ja niihin kohdistuvien odotusten muodostumisessa on käyttäjän tekninen tietous chatbotin toiminnasta. Ilman tietoa chatbotin teknisestä toiminnasta, chatboteilta odotetaan hyvinkin ihmismäistä käytöstä ja ihmisen kaltaista vuorovaikutuksellisuutta (Chen & Wang, 2018; Luger & Sellen, 2016). Lugerin ja Sellenin (2016) mukaan ilman tarkempaa tietoa chatbotin toiminnasta ja toteutuksesta, ihmismäisiä ominaisuuksia saatetaan liittää chatbottiin melkein naiivilla tavalla. He huomasivat tutkimuksessaan, että toiminnaltaan hyvinkin yksinkertaiseen chatbottiin voidaan liittää omia odotuksia vastaavia ihmismäisiä piirteitä, kuten esimerkiksi sarkasmin käyttöä, jota chatbotin tekninen toteutus ei edes mahdollista. He tunnistivat myös, kuinka virhetilanteissa osa käyttäjistä etsi ongelmaa usein ennemmin itsestä kuin olettaisi chatbotin tehneen virheen tai toimivan puutteellisesti (Luger & Sellen, 2016). Portela ja Granell-Canut (2017), Chen ja Wang (2018) sekä Luger ja Sellen (2016) huomioivat tutkimuksissaan chatbotin tekniseen toteutukseen perehtyneiden lähestyvän chatbotteja järjestelmällisemmin toimintojen perusteella, jolloin he eivät odota vuorovaikutuksen olevan yhtä ihmismäistä. Tämän takia chatbottien teknistä taustaa tuntevat käyttäjät osaavat hahmottaa sen toiminnallisuuden rajoja selkeämmin ja siten pystyvät käyttämään chatbotteja luontevammin (Portela & Granell-Canut, 2017; Chen & Wang, 2018; Luger & Sellen, 2016). Huomion arvoista on kuitenkin se, että vähemmän teknistä tietoutta omaavat käyttäjät suhtautuvat chatbotteihin avoimemmin, kun taas teknistä tietoa omaavat ovat skeptisempiä chatbottien potentiaalini ja sovellutusten suhteen (Portela & Granell-Canut, 2017). Teknisen tietouden voidaan todeta helpottavan chatbottien käyttöä, mutta voidaanko sen tulkita johtuvan chatbotin synnyttämien sosiaalisten reaktioiden vähemmästä kokemisesta? Antropomorfismilla ja chatbotin ihmismäisyydellä on nähty olevan positiivisia vaikutuksia chatbottien käyttökokemukseen, mutta samalla chatbottien lähestyminen toiminnallisuuden kautta käytön kautta tehokkaampaa. Ehkä tässä vaiheessa chatbottien kehitystä, ihmismäisyydellä on suurempi vaikutus käyttökokemukseen, kuin käytettävyyteen.

5 TUTKIMUSMENETELMÄ JA TUTKIMUKSEN TO- TEUTUS

Tässä luvussa esitellään tutkimuksen tarkoitus ja tutkimuksen toteutustapa. Aluksi esitellään tutkimuksen tavoite, tutkimuskysymykset ja niiden pohjalta luodut hypoteesit. Tämän jälkeen käsitellään tutkimuksen toteutusta perustelemalla aineistonkeruutapa, käydään läpi sen suunnittelu ja läpi vienti. Viimeisenä kerrotaan vielä, miten aineistonkeruussa kerätty aineisto analysoitiin tuloksien saamiseksi.

5.1 Tutkimuksen tavoite

Tutkimuksen tavoitteena on selvittää miten yksilöllisesti käyttäjät tulkitsevat chatbottien synnyttämiä sosiaalisia reaktioita ja miten chatbottien sisältämät sosiaaliset vinkit vaikuttavat tähän? Sosiaalisilla reaktioilla tarkoitetaan tässä yhteydessä chatbottien sisältämistä sosiaalisista vinkeistä syntyviä sosiaalisia reaktioita, joiden perusteella käyttäjät liittävät inhimillisiä ominaisuuksia chatbotteihin. Tutkimusongelmaa lähestytään seuraavien tutkimuskysymysten kautta:

- Ovatko chatbottien synnyttämät sosiaaliset reaktiot yhteneviä käyttäjien välillä?
- Vaikuttaako chatbottien sisältämien sosiaalisten vinkkien määrä chatbottien synnyttämiin sosiaalisiin reaktioihin käyttäjissä?

Ymmärtämällä chatbotista syntyvän mielikuvan muodostumista ihmisen ja chatbotin vuorovaikutuksessa, voidaan tätä tietoa hyödyntää tulevaisuudessa chatbottien suunnittelun apuna, ohjaamaan niiden toteutusta haluttuun käyttökontekstiin sopivaksi. Tutkimusongelmaa lähestytään Feinen ym. (2019) teoriaan perustuen, jonka mukaan chatbottien synnyttämät sosiaaliset reaktiot syntyvät chatbottien sisältämistä sosiaalisista vinkeistä. Tutkimuksen tulosten on

tarkoitus vahvistaa tätä aiempaa teoriaa ja tarjota siten syvempää tietämystä chatbottien synnyttämistä sosiaalisista reaktioista. Tutkimusasettelun pohjalta tutkimusongelmien ratkaisemiseksi on muodostettu seuraava testattava nollahypoteesi:

- H_0 : Chatbottien synnyttämissä sosiaalisissa reaktioissa ei ole eroa käyttäjien välillä

Nollahypoteesin vastahypoteesiksi on muodostettu seuraava hypoteesi:

- H_1 : Chatbottien synnyttämissä sosiaalisissa reaktioissa on eroa käyttäjien välillä

Hypoteesit H_0 ja H_1 perustuvat kirjallisuuskatsauksessa löydettyihin useisiin tutkimuksiin, joiden mukaan teknologioihin kohdistuvissa odotuksissa ja teknologioiden käytössä on käyttäjien välillä yksilöllisiä eroja (Gnewuch ym., 2018; Portela & Granell-Canut, 2017; Venkatesh ym., 2012; Danielescu & Christian, 2018; Dryer, 1999; Lee, Peng, Jin & Yan, 2006; Payne ym., 2013; Chen & Wang, 2018; Luger & Sellen, 2016). Hypoteesien avulla testataan pätevätkö nämä löydökset myös ihmisen ja chatbotin välisessä vuorovaikutuksessa.

Nollahypoteesin ja sen vastahypoteesin lisäksi muodostettiin myös kolmas hypoteesi pohjautuen tutkimuksiin, joiden mukaan teknologioiden sisältämällä sosiaalisilla vinkeillä on vaikutusta teknologioiden tulkintaan (Feine ym., 2019; Dryer, 1999; Walther, 2008; Koh & Sundar, 2010; Chaves & Gerosa, 2020; Sundar ym., 2016; Warwick & Shah, 2016). Tutkimuksissa on todettu myös chatbotin interaktiivisuudella olevan vaikutusta sen synnyttämiin sosiaalisiin reaktioihin (Go & Sundar, 2019; Portela & Granell-Canut, 2017; Jain ym., 2018). Myöskin chatbotin realistisuuden ja visuaalisuuden on tutkittu vaikuttavan sosiaalisten reaktioiden vahvuuteen kasvattavasti (Gong, 2008; Go & Sundar, 2019; Feine ym. 2019; Araujo, 2018; Rincón-Nigro & Deng, 2013; Beun ym., 2003; Ciechanowski ym., 2018). Näiden tutkimustulosten pohjalta muodostetaan vielä kolmas testattava hypoteesi:

- H_2 : Chatbottien synnyttämien sosiaalisten reaktioiden voimakkuus kasvaa suhteessa chatbotin sisältämien sosiaalisten vinkkien määrään

5.2 Tutkimusmenetelmän valinta

Tutkimusongelman ja tutkimuskysymysten pohjalta tutkimusmenetelmäksi valikoitui määrällinen kyselytutkimus verkkokyselynä toteutettuna. Vilkka (2007, s. 14) määrittelee määrällisen tutkimuksen tutkimustavaksi, jossa numeerisesta tiedosta tehdään tulkintoja ja havaintoja. Määrällistä tutkimusta käytetään selittämään, kuvaamaan, vertailemaan tai ennustamaan eri asioita tai ominaisuuksia (Vilkka, 2007, s. 19). Kysely on sopiva aineistonkeruumenetelmä

tilanteisiin, joissa halutaan tutkia yksittäisiä henkilöitä ja heidän mielipiteitään, asenteitaan, käyttäytymistään tai ominaisuuksiaan (Vilka, 2007, s. 28). Verkkokysely on siten perusteltu valinta tämän tutkimuksen aineistonkeruuseen, sillä tässä tutkimuksessa ollaan kiinnostuneita chatboteista yksittäisille käyttäjille syntyvistä mielikuvista ja ajatuksista, sekä niiden mahdollisista eroista. Pelkän määrällisen aineiston tueksi, aineistonkeruussa hyödynnetään monimetelmä-metodia keräämällä laadullista aineistoa avointen kysymyksen muodossa (Johnson & Onwuegbuzie, 2004). Avoimilla kysymyksillä on mahdollista kerätä vastaajien spontaanisti nousseita mielipiteitä (Vilka, 2007, s. 68). Näiden avointen kysymysten avulla pyritään saamaan syvempää ymmärrystä määrällisen aineiston tuloksien tulkinnan avuksi.

Tutkimuksen perusjoukkoon kuuluvat kaikki chatbottien potentiaaliset käyttäjät. Perusjoukolla tarkoitetaan henkilöistä, joita kyseenomaisessa tutkimuksessa halutaan tutkia (Vehkalahti, 2019, s. 43). Aineistonkeruun kannalta tämä tarkoittaa sitä, että tutkimukseen voi osallistua kuka tahansa henkilö, kennellä on mahdollisuus toimia chatbottien kanssa suoraan tai esimerkiksi jonkin palvelun kautta. Vilka (2007, s. 64) kuitenkin huomauttaa sähköisen kyselyn rajaavan vastaajista pois henkilöt, joilla ei ole mahdollisuutta käyttää tietotekniikkaa. Tämä ei kuitenkaan vaikuta juurikaan tämän tutkimuksen aineistonkeruuseen, sillä chatbottienkaan käyttö ei onnistu, mikäli henkilö ei pysty toimimaan tietotekniikan parissa. Aineiston tulisi olla otoskooltaan vähintään 100 havaintoyksikköä, jotta tutkimuksessa voidaan käyttää tilastollisia menetelmiä (Vilka, 2007, s. 17). Vehkalahti (2019, s. 43) korostaa kuitenkin otantaasetelman tärkeyttä ja miten se voi mahdollistaa luotettavien johtopäätösten tekemisen pienelläkin otoskoollla. Aineistonkeruun tavoitteena on kuitenkin kerätä vähintään 100 vastausta.

5.3 Tutkimuksen toteutus

Tutkimuksessa käytettävä verkkokysely vaatii verrattavissa olevat kysymykset erilaisille chatboteille, jotta tutkimukseen osallistujien vastauksien eroja on mahdollista vertailla. Tutkimus vaatii myös mittarin, jolla chatbottien synnyttämiä sosiaalisia reaktioita voidaan mitata. Lisäksi itse verkkokyselydokumentti pitää muodostaa ja julkaista vastauksien keräämiseksi ja myöhemmin analysoitavaksi. Näitä aiheita käsitellään seuraavissa aliluvuissa.

5.3.1 Chatbot-keskustelutilanteiden luonti

Tutkimusta varten käytiin läpi internetistä löytyviä chatbotteja, joita arvioitiin käymällä keskustelua niiden kanssa ja samalla analysoimalla niissä esiintyviä sosiaalisia vinkkejä Feinen ym. (2019) teoriaan perustuen. Chatbottien käytön ja vertailujen tuloksena, tutkimukseen valikoitiin seuraavat neljä chatbottia: Eliza (Eliza, 2021), Alice, (Alice, 2021), Mitsuku (Mitsuku, 2021) ja EvieBot (EvieBot,

2021). Valintojen perusteluna on chatbottien keskusteluun keskittyvä käyttötarkoitus, sekä niissä esiintyvien sosiaalisten vinkkien jakautuminen. Elizan ja Alicen havaittiin sisältävän suhteessa vähemmän sosiaalisia vinkkejä, kuin Mitsukun ja EvieBotin. Valintoihin vaikutti myös se, että jokaisen chatbotin kanssa vuorovaikutetaan samalla tavalla, eli tässä tapauksessa tekstin välityksellä.

Chatbottien valinnoissa keskityttiin juuri keskusteluun tarkoitettuihin chatbotteihin, jotta chatbotit olisivat toiminnallisuudeltaan verrannollisia keskenään. Tehtäväkeskeisten chatbottien kanssa vuorovaikuttaessa dialogi on usein toissijaista ja tärkeämpänä nähdään tietyn tehtävän suorittaminen onnistuneesti. Keskustelullisemmat chatbotit ovat tähän tutkimukseen otollisempi valinta, sillä niiden kanssa käyty vuorovaikutus ei ole yhtä suorituskeskeistä. Keskusteluun tarkoitetuissa chatboteissa keskiössä on käyttäjän kanssa vuorovaikuttaminen, joka luo hyvän perustan niiden synnyttämien sosiaalisten reaktioiden tutkimiselle.

Verkkokyselyssä chatboteista tehtävien tulkintojen verrattavuuden varmistamiseksi, tutkimus päätettiin toteuttaa hyödyntäen staattisia kuvia chatbottien kanssa käydyistä keskusteluista. Tällä tavalla yritetään luoda mahdollisimman samanlaiset lähtökohdat sosiaalisten reaktioiden synnylle, jolloin käyttäjien kokemien sosiaalisten reaktioiden mahdollisissa eroissa selittävänä muuttujana ovat käyttäjät itse. Staattisia kuvia chatbottien kanssa käydyistä keskusteluista on hyödynnetty aiemminkin chatbottien tutkimuksessa (ks. Chung ym., 2018).

Tutkimuksessa käytettävien chatbottien keskustelukuvien luomiseksi, muodostettiin yksinkertainen skenaario ohjaamaan jokaisen valitun chatbotin kanssa käytyä keskustelua. Skenaarion keskustelunaiheeksi valittiin emotionaalisen tuen hakeminen chatbotilta, jonka on havaittu olevan yksi yleisistä chatbotteihin kohdistuvista odotuksista (Zamora, 2017). Luodussa skenaariossa käyttäjä kertoo chatbotille kokevansa stressiä. Skenaario esitettiin chatboteille seuraavanlaisen dialogikulun kautta: tervehdys ja esittäytyminen, kuulumisten kysyminen, stressin tuntemisen esittäminen ja keskustelun lopettaminen. Dialogikulkua seurattiin suuntaa antavasti, eli chatbotin kanssa käydyn dialogin kulkiessa hyvin, sitä jatkettiin luonnolliseen tapaan, mutta keskustelun pysähtyessä siirryttiin esitellyn dialogikulun seuraavaan kohtaan.

Chatbottien kanssa käydyistä keskusteluista luodut kuvat ovat näkyvillä valmiissa kyselylomakkeessa liitteessä 2. Luodut kuvat analysoitiin Feinen ym. (2019) teoriaan pohjautuen ja keskusteluista tunnistetut ja ilmenevät sosiaaliset vinkit esitellään taulukossa 3. Tunnistetut sosiaalisten vinkkien kategoriat painottuivat selkeästi verbaalisiin ja visuaalisiin vinkkeihin, joskin Elizan kohdalla havaittiin myös yksi näkymättömiin vinkkeihin kuuluva sosiaalinen vinkki. Elizan kohdalla sosiaalisia vinkkejä tunnistettiin yhteensä 9, Alicella 8, Mitsukulla 21 ja EvieBotilla 20.

TAULUKKO 3 Tutkimuksen chatboteista tunnistetut sosiaaliset vinkit.

<i>Chatbot ja sen sisältämät sosiaaliset vinkit</i>		<i>Eliza</i>	<i>Alice</i>	<i>Mitsuku</i>	<i>EvieBot</i>
<i>Verbaalinen</i>	Sisältö	Tervehdys, Small talk, Menneeseen viittaaminen, Kysymykset itsestä,	Tervehdys, Small talk, Menneeseen viittaaminen, Kiittäminen, Itsestä kertominen,	Tervehdys, Keskustelun ohjaus, Small talk, Menneeseen viittaaminen, Kysymykset itsestä, Myötämielisyys, Kiittäminen	Tervehdys, Myötämielisyys, Keskustelun ohjaus, Small talk, Ylistys, Itsestä kertominen, Myötämielisyys, Vinkit ja neuvot
	Tyyli	Virkkeiden monimutkaisuus, Kielen painokkuus	Virkkeiden monimutkaisuus	Virkkeiden monimutkaisuus	Virkkeiden monimutkaisuus
<i>Visuaalinen</i>	Tietokonevälitteinen viestintä	Fontti	Fontti	Emoji/gif, Fontti	Fontti
	Agentin ulkoasu	Nimi	Nimi	Nimi, Sukupuoli, Realistisuus, Kasvonpiirteet, Fotorealismi, Ikä, Vaatetus, Väryty, 2D/3D-visualisointi	Nimi, Sukupuoli, Realistisuus, Kasvonpiirteet, Fotorealismi, Ikä, Väryty, 2D/3D-visualisointi
	Proksemiikka	-	-	Taustakuva, Vuorovaikutuksellinen etäisyys	Taustakuva, Vuorovaikutuksellinen etäisyys
<i>Näkyvät</i>	Kronemiikka	Aloitteen tekeminen	-	-	-
	<i>Yhteensä (kpl)</i>	9	8	21	20

5.3.2 Sosiaalisten reaktioiden mittaaminen

Chatbotin ihmisessä aiheuttamien sosiaalisten reaktioiden selvittämiseksi, kyselyä varten muodostettiin sitä varten mittari. Vehkalahti (2019, s. 12) määrittelee

kyselytutkimuksessa käytettävän mittarin olevan kysymysten tai väittämien kokoelma, joilla pyritään mittaamaan erilaisia ilmiöitä. Tutkimuksessa käytettävän mittarin voi luoda itse tai hyödyntää aiemmin muodostettua mittaria (Vehkalahti, 2019, s. 12). Ihmisen ja chatbotin välisen vuorovaikutuksen mittaamiseen ei ole olemassa standardia tai muuta vakiintunutta mittaria. Sen sijaan ihmisen ja robotin väliseen vuorovaikutukseen on kehitetty muutamia mittareita, joista yksi käytetyimmistä on Bartneekin, Croftin ja Kulicin (2008) kehittämä Godspeed-kysely. Godspeed-kyselyssä ihmisen ja robotin välistä vuorovaikutusta mitataan viiden eri ulottuvuuden kautta: antropomorfismin, elollisuuden, miellyttävyyden, havaitun älykkyyden ja turvallisuudentunteen. Nämä ulottuvuudet koostuvat useista Osgoodin asteikolla olevista adjektiivipareista (Bartneck ym., 2008). Osgoodin asteikko on niin sanottu asenneasteikko ja se soveltuu asenteiden ja mielipiteiden mittaamiseen (Vilkkä, 2007, s. 45).

Vaikka Godspeed-kysely on suunnattu ja suunniteltu ihmisen ja robotin välisen vuorovaikutuksen arviointiin, on siinä silti yhtymäkohtia myös ihmisen ja chatbotin väliseen vuorovaikutukseen. Chatbotin ja robotin erona voidaan pitää sitä, että robotilla on vuorovaikutustoiminnallisuuden lisäksi fyysinen olemus. Tästä syystä Godspeed-kyselyn mittareita voidaan soveltaa myös chatbotteihin, erottamalla siitä robotin fyysistä ulottuvuutta mittaavat kohdat. Näitä fyysistä ulottuvuutta mittaavia kohtia ovat elollisuuden ja turvallisuudentunteen mittarit, joiden sisältämät adjektiiviparit kohdistuvat vahvasti robotin eleisiin ja elehdintään. Jäljelle jäävät ulottuvuudet, antropomorfismi, miellyttävyys ja havaittu älykkyys, koskevat sen sijaan ihmisen ja robotin välisen vuorovaikutuksen psyykkistä ulottuvuutta. Luvussa 4 tehtyjen huomioiden chatbottien synnyttämistä sosiaalisista reaktioista voidaan havaita sisältyvän antropomorfismin, miellyttävyyden ja havaitun älykkyyden -kategorioihin. Tämän havainnon perusteella, Godspeed-kyselyn sisältämät mittarien nähdään sopivia tämän tutkimuksen tarkoituksiin.

Godspeed-kyselyä mukautettiin tutkimukseen sopivaksi poistamalla kyselystä edellä mainitut elollisuuden ja turvallisuudentunteen kohdat. Tämän lisäksi, poistettiin havaitun älykkyyden mittarista yksi tehtäväorientoitunut adjektiivipari, joka ei sopinut keskustelevien chatbottien arviointiin. Kyselyssä käytettävien staattisten keskustelukuvien takia, kyselyn antropomorfismin ulottuvuudesta poistettiin myös liikkeen sujuvuuteen viittaava adjektiivipari. Lopullinen tutkimuksessa käytettävä mukautus Godspeed-kyselystä on esitetty taulukossa 4.

TAULUKKO 4 Mukautettu ja suomennettu Godspeed-kysely (Bartneck ym., 2008)

<i>Mitattava ulottuvuus</i>	<i>Lyhenne</i>	<i>Suomennetut adjektiiviparit (alkuperäinen suluissa)</i>
<i>Antropomorfismi</i>	ANT1	Teennäinen (Fake) – Luonnollinen (Natural)
	ANT2	Konemainen (Machinelike) – Ihmismäinen (Humanlike)
	ANT3	Tiedoton (Unconscious) – Tietoinen (Conscious)
	ANT4	Keinotekoinen (Artificial) – Elävä (Lifelike)
<i>Miellyttävyys</i>	MIE1	Inhottava (Dislike) – Pidettävä (Like)
	MIE2	Epäystävällinen (Unfriendly) – Ystävällinen (Friendly)
	MIE3	Tyly (Unkind) – Kiltti (Kind)
	MIE4	Epämiellyttävä (Unpleasant) – Miellyttävä (Pleasant)
	MIE5	Kamala (Awful) – Mukava (Nice)
<i>Havaittu älykyys</i>	HAV1	Tietämätön (Ignorant) – Tietäväinen (Knowledgeable)
	HAV2	Vastuuton (Irresponsible) – Vastuullinen (Responsible)
	HAV3	Tyhmä (Unintelligent) – Älykäs (Intelligent)
	HAV4	Vaikeaselkoinen (Foolish) – Ymmärrettävä (Sensible)

5.3.3 Kyselyn toteutus

Muodostettua kyselyä testattiin ennen sen julkaisua tutkimuksen ulkopuolisen henkilön toimesta, jolla varmistuttiin kyselyn ymmärrettävyydestä ja rakenteesta. Kysely toteutettiin Webropol-palvelun kautta verkkokyselynä 16.3.2021-30.2.2021 välisenä aikana. Kyselyä levitettiin hyödyntämällä Jyväskylän yliopiston sähköpostilistaa, Jyväskylän yliopiston ainejärjestöjä, sekä tutkijan omia kontakteja. Kysely koostui yhteensä kuudesta osasta. Verkkokyselyn ensimmäisessä osassa kerättiin vastaajien demografisia tietoja, joiden avulla voidaan tarkastella mahdollisia taustatekijöitä tuloksista. Kysytyjä demografisia tietoja olivat syntymävuosi, sukupuoli, tietoteknisen osaamisen taso sekä olivatko vastaaja käyttänyt chatbotteja aiemmin. Tämän jälkeen kyselyn toisessa osassa esitettiin lyhyt opastus tutkimuksessa käytettävien keskustelukuvien läpikäymiseen ja kysymyksiin vastaamiseen. Loput neljä osaa sisälsivät luodut chatbottien keskustelukuvat, jotka esitettiin jokainen yksi kerrallaan omassa osassaan. Jokaisessa näistä osista oli chatbotin kanssa käydystä keskustelusta luotu kuva ja sen jälkeen mukautettu Godspeed-kysely. Tämän osiossa on myös vapaaehtoinen avoin tekstilaatikko, johon vastaaja voi omin sanoin kertoa chatbotista nousseita mielteitään. Kyselyn lopuksi esitettiin kiitos kyselyyn vastaamisesta. Kysely on näkyvissä kokonaisuudessaan liitteessä 2.

Verkkokyselyssä esitettävän Godspeed-kyselyn esitysmuotoa muokattiin hieman, kysymysten asettelusta tai järjestyksestä vastaajalle syntyvien oletuksien tai tulkittujen asiayhteyksien minimoimiseksi. Mukautetun Godspeed-kyselyn sisältämät 13 adjektiiviparia esitettiin satunnaisessa järjestyksessä, jolloin samaan kategoriaan kuuluvat adjektiiviparit eivät olisi kyselyssä peräkkäin. Tämän lisäksi viisi kohtaa esitettiin käänteisesti vaihtamalla niiden sisältämien adjektiivien paikkaa keskenään. Tällä tavalla eri kategorioiden skaalojen ääriarvot eivät ole aina samalla puolella itse kyselyssä.

5.4 Aineiston analysointi

Verkkokyselyllä kerätty aineisto ladattiin Webropol-palvelusta .sav-tiedostomuodossa, jonka jälkeen sen käsittely aloitettiin SPSS-ohjelmalla. Ensimmäiseksi aineiston muuttujat nimettiin kuvaavimmilla nimillä käsittelyn helpottamiseksi. Tämän jälkeen aineisto käytiin läpi puuttuvien tai virheellisten arvojen varalta. Puuttuvia arvoja löytyi odotetusti ainoastaan kyselyn vapaaehtoisista avoimista kysymyksistä, muiden kohtien ollessa pakollisia täyttää. Muutamaan avoimeen kysymykseen oli vastattu pelkällä ”-”-merkillä, jotka tulkittiin tyhjiksi vastauksiksi ja muutettiin siten puuttuviksi arvoiksi.

Vastausten oikeellisuuden tarkistamisen jälkeen, aloitettiin aineiston muokkaaminen analyysia varten. Aluksi verkkokyselyä varten käännetyt mukautetun Godspeed-kyselyn kohdat käännettiin takaisin alkuperäiseen muotoonsa, jotta adjektiiviparien skaalat ovat samoin päin keskenään. Tämän jälkeen syntymävuosista laskettiin vastaajien iät uuteen muuttujaan, jotka vielä edelleen lajiteltiin 10 ikävuoden ryhminä kolmanteen muuttujaan.

Aineiston muokkaamisen jälkeen luotiin mukautetun Godspeed-kyselyn kategorioista summamuuttujat: antropomorfismi, miellyttävyys ja havaittu älykkyys. Luodut summamuuttujat skaalattiin lisäksi takaisin alkuperäiselle skaalalle (1-5), niiden vertailun helpottamiseksi. Godspeed-kyselyn pohjalta luotujen summamuuttujien reliabiliteetti tarkistettiin laskemalla siitä Cronbachin alfa, joka mittaa kuinka hyvin valitut muuttujat mittaavat samaa asiaa (Metsämuuronen, 2011, s. 545). Summamuuttujien normaalijakautuneisuus tarkistettiin graafisesti niistä tehtyjen histogrammien kautta (liite 1). Summamuuttujien tunnuslukuja tarkisteltiin myös epämääräisten arvojen varalta.

Summamuuttujien analysointimenetelmänä käytettiin yksisuuntaista varianssianalyysiä, joka sopii eri luokkien ja ryhmien keskiarvojen vertailemiseen (Vehkalahti, 2019, s. 139). Menetelmä sopii siten tämän aineiston analysointiin, jossa halutaan vertailla summamuuttujien arvoja eri chatbottien välillä. Vilkan (2007, s. 57) mukaan eri ryhmien välisen vertailun vaatimuksena on, että joka ryhmässä on vähintään 30 havaintoyksikköä. Tämä vaatimus toteutuu tutkimuksessa summamuuttujien sisältäessä 100 havaintoyksikköä jokaista chatbottia kohden.

Varianssianalyysin käyttöön liittyy kolme oletusta: havaintojen tulee olla toisistaan riippumattomia, otoskeskiarvojen tulee olla riittävän normaalijakautuneet ja viimeisenä ryhmien varianssien tulee olla yhtä suuret (Metsämuuronen, 2011, s. 784). Vastausten riippumattomuus toteutui tutkimuksessa aineistonkeruumenetelmän ansiosta, joka ei ollut aikaan tai paikkaan sidonnainen. Muuttujien normaalijakautuneisuuden voi tarkistaa silmämääräisesti grafiikan perusteella (Metsämuuronen, 2011, s. 785). Summamuuttujien histogrammeja tarkastelemalla, muuttujien todettiin olevan tarpeeksi normaalijakautuneita. Yksisuuntainen varianssianalyysi ei ole kovin herkkä aineiston normaalijakautuneisuuden virheille (Blanca, Alarcón, Arnau, Bono & Bendayan, 2017).

Oletus varianssien yhtäsuuruudesta varmistettiin Levenen testillä. Mikäli varianssianalyysin oletukset eivät toteutuneet, käytettiin sen sijasta vastaavaa, mutta ei-parametristä Kruskal-Wallis-testiä (Metsämuuronen, 2011, s. 785).

Varianssianalyysin oletukset eivät toteutuneet demografisten tekijöiden kohdalla, jossa havaintoyksiköiden minimimäärä 30 ei täytynyt jokaisen vertailtavan ryhmän kohdalla. Demografisten tekijöiden analysoinnissa käytettiin siksi Kruskal-Wallis-testiä, joka soveltuu pienemmille otoskoille ja jos normaali-jakautuneisuudesta ei olla varmoja (Metsämuuronen, 2011, s. 785).

Avointen kysymysten analysointi tehtiin teemoittelun avulla. Saadut vastaukset käytiin läpi huolellisesti ja samankaltaiset huomiot luokiteltiin omiksi teemoikseen. Vastauksia tulkittiin varsinkin chatbotin antropomorfismiin, miellyttävyyteen ja havaittuun älykkyyteen liittyen, mutta analysoinnissa huomioitiin myös muihin aiheisiin liittyvät kommentit.

6 TUTKIMUKSEN TULOKSET

Tässä luvussa käsitellään aineistonkeruun tuloksia ja niistä tehtyjä havaintoja. Aluksi aineisto esitellään yleisesti ja kuvaillaan tutkimukseen vastaajien demografisia tekijöitä. Tämän jälkeen käydään läpi aineistosta muodostettujen summamuuttujien arvoja ja niille suoritettuja analyyskejä ja testejä. Lopuksi käydään vielä läpi vastaajien avoimiin kysymyksiin antamia vastauksia ja niistä nousevia huomioita.

6.1 Aineiston kuvailu

Aineistonkeruun tuloksena saatiin kerättyä 100 havaintoyksikön kokoinen aineisto, joka on juuri riittävä tilastollisten menetelmien käyttöön. Taulukossa 5 on esitelty kyselyyn vastanneiden demografisia tietoja. Aineiston vastaajista 41 % oli miehiä ja 53 % naisia. Vastanneista 3 % ilmoitti sukupuolekseen ”muu” ja loput 3 % eivät halunneet mainita sukupuoltaan. Tutkimukseen vastanneiden keski-ikä oli 29,64 vuotta. Nuorin tutkimukseen osallistunut oli 16-vuotias ja vanhin 59-vuotias. Enemmistö vastaajista olivat kuitenkin 25–34-vuotiaita, prosenttiosuudella 52. Seuraavaksi suurin ikärymä oli 16–24-vuotiaat, joihin kuului 29 % vastanneista. 35–44-vuotiaiden osuus oli 10 %, 45–55-vuotiaiden 6 % ja yli 55-vuotiaiden 3 %. Tutkimukseen vastanneiden iän voidaan todeta painottuvan nuoriin aikuisiin.

Vastaajien tietotekninen osaaminen jakautui kolmeen kategoriaan viidestä. Enemmistö vastaajista määritteli tietoteknisen osaamisensa ”Harrastuneisuutta”-kategoriaan prosenttiosuudella 43. Ammattilaiseksi itsensä määrittelivät 26 % vastanneista ja pelkät perustaidot omaavia olivat loput 31 %. ”Ei ollenkaan” ja ”Välttämättömät taidot” -kategorioihin ei kuulunut yksikään vastaaja, joka ei ole yllättävää, sillä verkkokysely rajaa tietotekniikkaa käyttämättömät pois vastaajista. Aiempaa kokemusta chatboteista oli 86 % vastanneista ja vastavuoroisesti 14 % ilmoitti, ettei heillä ollut aiempaa kokemusta chatboteista.

TAULUKKO 5 Kyselyn vastaajien demografiset tiedot

<i>Muuttuja</i>		<i>N</i>	<i>%</i>
<i>Sukupuoli</i>	Mies	41	41
	Nainen	53	53
	Muu	3	3
	En halua sanoa	3	3
<i>Ikä</i>	16–24	29	29
	25–34	52	52
	35–44	10	10
	45–54	6	6
	yli 55	3	3
<i>Tietotekninen osaaminen</i>	Perustaidot	31	31
	Harrastuneisuutta	43	43
	Ammattilainen	26	26
<i>Aiempaa kokemusta chatboteista</i>	Kyllä	86	86
	Ei	14	14
	<i>Yhteensä</i>	100	100

6.2 Mukautetun Godspeed-kyselyn tulokset

Aineistonkeruussa käytetyn kyselyn perustuessa Bartneckin ym. (2008) luomaan Godspeed-kyselyyn, muuttujien ryhmittely ja niistä muodostettavat summamuuttujat olivat jo ennalta tiedossa. Muodostettavien summamuuttujien reliabiliteetti tämän tutkimuksen aineistossa tarkistettiin kuitenkin vielä konfirmatorisesti Cronbachin alfa -testillä. Summamuuttujien voidaan tulkita olevan oikeellisia, jos Cronbachin alfa saa arvon $\alpha > 0,60$ (Metsämuuronen, 2011, s. 549). Cronbachin alfa -testin tulokset on kerätty taulukkoon 6. Muodostettujen summamuuttujien saamat alfa-arvot ovat jokaisessa kohdassa yli 0,60, joten summamuuttujien todetaan läpäisevän Cronbachin alfa -testin. Testin yhteydessä tarkistettiin myös Cronbachin alfa -testin yhteydessä luodut korrelatiomatriisit muuttujien välillä. Näissä ei havaittu mitään epänormaalia, vaan summamuuttujien sisältämät muuttujat korreloivat odotetusti keskenään. Näiden tulosten perusteella summamuuttujat todettiin oikeellisiksi.

TAULUKKO 6 Mukautetun Godspeed-kyselyn summamuuttujien reliabiliteetti

<i>Chatbot</i>	<i>Summamuuttuja</i>	<i>Muuttujat</i>	<i>Cronbachin a</i>
<i>Eliza</i>	Antropomorfismi	ANT1, ANT2, ANT3, ANT4	0,700
	Miellyttävyys	MIE1, MIE2, MIE3, MIE4, MIE5	0,841
	Havaittu älykkyys	HAV1, HAV2, HAV3, HAV4	0,628
<i>Alice</i>	Antropomorfismi	ANT1, ANT2, ANT3, ANT4	0,882
	Miellyttävyys	MIE1, MIE2, MIE3, MIE4, MIE5	0,914
	Havaittu älykkyys	HAV1, HAV2, HAV3, HAV4	0,770
<i>Mitsuku</i>	Antropomorfismi	ANT1, ANT2, ANT3, ANT4	0,865
	Miellyttävyys	MIE1, MIE2, MIE3, MIE4, MIE5	0,888
	Havaittu älykkyys	HAV1, HAV2, HAV3, HAV4	0,713
<i>EvieBot</i>	Antropomorfismi	ANT1, ANT2, ANT3, ANT4	0,850
	Miellyttävyys	MIE1, MIE2, MIE3, MIE4, MIE5	0,901
	Havaittu älykkyys	HAV1, HAV2, HAV3, HAV4	0,768

Summamuuttujien tunnusluvut on esitetty taulukossa 7. Tunnuslukujen perusteella voidaan havaita summamuuttujissa olevan eroja eri chatbottien välillä. Varsinkin Elizan kohdalla antropomorfismin, miellyttävyys ja havaitun älykkyys nähdään olevan selkeästi matalampia verrattuna muihin chatbotteihin. Alicen, Mitsukun ja EvieBotin tunnuslukujen välillä ero ei ole yhtä enää selkeää, joskin arvoissa nähdään pieni nousu verratessa Alicen arvoa Mitsukun arvoihin ja Mitsukun arvoja edelleen EvieBotin arvoihin.

TAULUKKO 7 Chatbottien summamuuttujien tunnusluvut

<i>Chatbot</i>	<i>Summamuuttuja</i>	<i>Keskiarvo</i>	<i>Keskihajonta</i>	<i>Moodi</i>	<i>Mediaani</i>
<i>Eliza</i>	Antropomorfismi	1,80	0,66	1,00	1,75
	Miellyttävyys	1,97	0,64	2,40	2,00
	Havaittu älykkyys	2,18	0,64	1,75	2,25
<i>Alice</i>	Antropomorfismi	2,68	0,88	2,50	2,75
	Miellyttävyys	3,56	0,76	4,00	3,60
	Havaittu älykkyys	2,96	0,70	2,75	3,00
<i>Mitsuku</i>	Antropomorfismi	2,95	0,91	2,75	3,00
	Miellyttävyys	3,78	0,63	4,00	3,90
	Havaittu älykkyys	3,36	0,68	3,25	3,38
<i>EvieBot</i>	Antropomorfismi	3,12	0,92	3,00	3,00
	Miellyttävyys	3,97	0,71	4,00	4,00
	Havaittu älykkyys	3,60	0,69	3,75	3,75

Varianssianalyysiä varten summamuuttujien varianssien yhtäsuuruus testattiin Levenen testillä, jonka tulokset on esitetty taulukossa 8. Testin nollahypoteesina on, että varianssit eivät ole yhtä suuret. Testin tuloksista havaitaan, että miellyttävyys ja havaitun älykkyys kohdalla arvojen varianssit ovat yhtä suuret ($p > 0,05$). Sen sijaan antropomorfismin kohdalla testi ei mene läpi ($p < 0,05$). Näiden tulosten perusteella miellyttävyys ja havaitun älykkyys kohdalla yksisuuntainen varianssianalyysi on sopiva analyysimenetelmä, mutta antropomorfismin kohdalla tulee käyttää Kruskal-Wallis-testiä.

TAULUKKO 8 Levenen testin tulokset

<i>Muuttuja</i>	<i>Ryhmä</i>	<i>Levenen tunnusluku</i>	<i>df1</i>	<i>df2</i>	<i>p-arvo</i>
<i>Antropomorfismi</i>	Chatbotit	3,308	3	396	0,020
<i>Miellyttävyyys</i>	Chatbotit	0,246	3	396	0,866
<i>Havaittu älykkyys</i>	Chatbotit	0,137	3	396	0,938

Antropomorfismin arvojen vertailemiseksi chatbottien välillä suoritettiin Kruskal-Wallis-testi, jonka tulokset on esitetty taulukossa 9. Kruskal-Wallis-testin nollahypoteesina on, että muuttujan keskiarvojen välillä on eroa ryhmien välillä. Testin tuloksesta havaitaan nollahypoteesin kumoutuvan ja antropomorfismin keskiarvojen eroavan tilastollisesti erittäin merkitsevästi ($p < 0,001$) eri chatbottien välillä.

TAULUKKO 9 Kruskal-Wallis-testin tulokset antropomorfismin ja eri chatbottien välillä

<i>Muuttuja</i>	<i>Selittävä ryhmä</i>	<i>Tunnusluku</i>	<i>df</i>	<i>p-arvo</i>
<i>Antropomorfismi</i>	Chatbot	109,064	3	0,000

Kruskal-Wallis-testi ei kuitenkaan kerro minkä ryhmien välillä antropomorfismin arvot eroavat toisistaan, jonka takia testiä jatkettiin suorittamalla vielä parivertailut chatbottien välillä. Parivertailun tulokset näkyvät taulukossa 10. Tuloksista havaitaan Elizan antropomorfismilla olevan tilastollisesti erittäin merkitsevä ero ($p < 0,001$) Alicen, Mitsukun ja EvieBotin saamiin arvoihin. Myöskin Alicen ja EvieBotin välillä on tilastollisesti melkein merkitsevä ero ($p < 0,05$). Sen sijaan Alicen ja Mitsukun ($p = 0,391$) sekä Mitsukun ja EvieBotin ($p = 1$) välillä antropomorfismissa ei ole tilastollisesti merkitsevää eroa.

TAULUKKO 10 Parivertailun tulokset antropomorfismin ja chatbottien välillä

<i>Muuttuja</i>	<i>Chatbotit</i>	<i>p-arvo^a</i>
<i>Antropomorfismi</i>	Eliza - Alice	0,000
	Eliza - Mitsuku	0,000
	Eliza - EvieBot	0,000
	Alice - Mitsuku	0,391
	Alice - EvieBot	0,023
	Mitsuku - EvieBot	1,000

a. Bonferroni-korjattu arvo

Miellyttävyyden ja havaitun älykkyuden summamuuttujat täyttivät varianssi-analyysin oletukset, joten niiden keskiarvojen erojen analysointiin käytettiin yksisuuntaista varianssi-analyysiä. Testin tulokset näkyvät taulukossa 11. Varianssi-analyysin nollahypoteesina on, että keskiarvojen välillä ei ole eroa ryhmien välillä. Testin tuloksista kuitenkin selviää, että miellyttävyyden sekä havaitun älykkyuden keskiarvoilla on tilastollisesti erittäin merkitsevä ($p < 0,001$) ero chatbottien välillä, eli nollahypoteesi kumoutuu. Eta²-arvo kertoo, että 58 % miellyttävyydestä ja 39 % havaitusta älykkyydestä selittyy chatbot valinnalla.

TAULUKKO 11 Miellyttävyyden ja havaitun älykkyyden yksisuuntaisten varianssianalyysien tulokset eri chatbottien suhteen

<i>Muuttuja</i>	<i>Selittävä ryhmä</i>	<i>F-testi</i>	<i>p-arvo</i>	<i>eta²</i>
<i>Miellyttävyys</i>	Chatbot	178,629	0,000	0,575
<i>Havaittu älykyys</i>	Chatbot	83,522	0,000	0,388

Ryhmäkohtaisten erojen tunnistamiseksi, varianssianalyysin jälkeen suoritettiin myöskin parivertailut eri ryhmien välillä. Parivertailun tulokset on kerätty taulukkoon 12. Tuloksista havaitaan, että Elizan kohdalla miellyttävyyden keskiarvot ovat tilastollisesti erittäin merkitseviä ($p < 0,001$) jokaisen muun chatbotin suhteen. Myöskin Alicen ja EvieBotin välillä miellyttävyys eroaa tilastollisesti erittäin merkitsevästi ($p < 0,001$). Sen sijaan Alicen ja Mitsukun ($p = 0,098$) sekä Mitsukun ja EvieBotin ($p = 0,223$) välillä ei miellyttävyydessä ei ole tilastollisesti merkitsevää eroa.

Havaitun älykkyyden keskiarvojen suhteen Elizan arvot eroavat myöskin tilastollisesti erittäin merkitsevästi ($p < 0,001$) jokaisesta muusta chatbotista. Myös Alicen kohdalla havaitun älykkyyden keskiarvot eroavat tilastollisesti erittäin merkitsevästi ($p < 0,001$) muista chatboteista. Ainoastaan Mitsukun ja EvieBotin välillä havaitussa älykkyydessä ei ole tilastollisesti merkittävää eroa ($p = 0,071$).

TAULUKKO 12 Varianssianalyysien parivertailun tulokset miellyttävyydelle ja havaitulle älykkyydelle chatbottien suhteen

<i>Selitettävä muuttuja</i>	<i>Chatbotit (I)</i>	<i>Chatbotit (J)</i>	<i>Keskiarvojen ero (I-J)</i>	<i>p-arvo^a</i>
<i>Miellyttävyys</i>	<i>Eliza</i>	Alice	-1,592***	0,000
		Mitsuku	-1,816***	0,000
		EvieBot	-2,002***	0,000
	<i>Alice</i>	Eliza	1,592***	0,000
		Mitsuku	-0,224	0,098
		EvieBot	-0,410***	0,000
	<i>Mitsuku</i>	Eliza	1,816***	0,000
		Alice	0,224	0,098
		EvieBot	-0,186	0,223
	<i>EvieBot</i>	Eliza	2,002***	0,000
		Alice	0,410***	0,000
		Mitsuku	0,186	0,223
<i>Havaittu älykyys</i>	<i>Eliza</i>	Alice	-0,783***	0,000
		Mitsuku	-1,183***	0,000
		EvieBot	-1,418***	0,000
	<i>Alice</i>	Eliza	0,783***	0,000
		Mitsuku	-0,400***	0,000
		EvieBot	-0,635***	0,000
	<i>Mitsuku</i>	Eliza	1,185***	0,000
		Alice	0,400***	0,000
		EvieBot	-0,235	0,071
	<i>EvieBot</i>	Eliza	1,418***	0,000
		Alice	0,635***	0,000
		Mitsuku	0,235	0,071

*** Tilastollisesti erittäin merkitsevä tasolla $p < 0,001$

a. Tukey-korjattu arvo

Luvussa 5.4 esiteltyjen perustelujen takia, demografisten tekijöiden mahdollisten vaikutusten tarkastelemisessa käytettiin ainoastaan Kruskal-Wallis-testiä. Tutkittavina demografisina tekijöinä olivat sukupuoli, ikäryhmä, tietotekninen osaaminen ja aiempi kokemus chatboteista. Kruskal-Wallis-testi suoritettiin jokaisen chatbotin kohdalla, vertaillen antropomorfismin, miellyttävyyden ja havaitun älykkyyden keskiarvoja demografisten tekijöiden ryhmiin. Kruskal-Wallis-testin tulokset näkyvät taulukossa 13.

Testin tuloksien perusteella demografisten tekijöiden aiheuttamia tilastollisesti merkitseviä eroja löytyy ainoastaan eri sukupuolien sekä ikäryhmien sisällä. Sukupuolella nähdään olevan tilastollisesti melkein merkitsevää ($p < 0,05$) eroa Mitsukun miellyttävyyden ja EvieBotin havaitun älykkyyden suhteen. Ikäryhmien kohdalla melkein tilastollisesti melkein merkitseviä ($p < 0,05$) eroja löytyi antropomorfismin kohdalla Elizan, Alicen ja EvieBotin suhteen ja havaitun älykkyyden kohdalla Alicen, Mitsukun ja EvieBotin suhteen. Ikäryhmällä huomataan olevan EvieBotin miellyttävyyden suhteen tilastollisesti merkitsevää ($p < 0,01$) ja Mitsukun miellyttävyyden suhteen tilastollisesti erittäin merkitsevää eroa ($p < 0,001$). Tietoteknisen osaamisen tai aiemman kokemuksen

chatboteista ei näiden tulosten perusteella nähdä vaikuttavan antropomorfismin, miellyttävyyden tai havaitun älykkyyden tulkintoihin.

TAULUKKO 13 Demografisten tekijöiden tilastolliset merkitsevyydet summamuuttujien suhteen eri chatboteille

Chatbot	Summamuuttuja	Sukupuoli	Ikäryhmä	Tietotekninen osaaminen	Kokemusta chatboteista
Eliza	Antropomorfismi	0,899	0,032*	0,878	0,884
	Miellyttävyys	0,507	0,125	0,986	0,175
	Havaittu älykkyyys	0,645	0,330	0,536	0,515
Alice	Antropomorfismi	0,058	0,040*	0,056	0,553
	Miellyttävyys	0,136	0,089	0,168	0,542
	Havaittu älykkyyys	0,564	0,013*	0,132	0,912
Mitsuku	Antropomorfismi	0,072	0,055	0,152	0,063
	Miellyttävyys	0,032*	0,000***	0,232	0,170
	Havaittu älykkyyys	0,106	0,040*	0,110	0,453
EvieBot	Antropomorfismi	0,120	0,037*	0,424	0,715
	Miellyttävyys	0,711	0,005**	0,532	0,833
	Havaittu älykkyyys	0,017*	0,012*	0,128	0,426

*** Tilastollisesti erittäin merkitsevä tasolla $p < 0,001$

** Tilastollisesti merkitsevä tasolla $p < 0,01$

* Tilastollisesti melkein merkitsevä tasolla $p < 0,05$

Ikäryhmien ryhmäkohtaisten erojen tunnistamiseksi suoritettiin jälleen parivertailut vertailtavien ryhmien välillä. Taulukossa 14 on esitetty parivertailun tilastollisesti merkitsevät ($p < 0,05$) tulokset. Elizan kohdalla ikäryhmien välillä ei ole minkäänlaista tilastollisesti merkittävää eroa. Sen sijaan Alicen kohdalla tilastollisesti melkein merkitsevää eroa löytyi antropomorfismin ja havaitun älykkyyden suhteen. Mitsukun kohdalla, tuloksista nähdään tilastollisesti merkitsevää eroa olevan pelkästään miellyttävyyden kohdalla. EvieBotin kanssa ikäryhmien välisiä eroja löytyi jokaisen summamuuttujan suhteen. Eroavat ikäryhmät näkyvät tarkemmin taulukosta 14.

TAULUKKO 14 Chatbottien summamuuttujien tilastollisesti merkitsevät erot ikäryhmien välillä

<i>Chatbot</i>	<i>Summamuuttuja</i>	<i>Eroavat ikäryhmät</i>	<i>p-arvo^a</i>
<i>Alice</i>	Antropomorfismi Havaittu älykkyys	35–44 v - yli 55 v	0,020
		35–44 v - yli 55 v	0,011
		45–54 v - yli 55 v	0,026
<i>Mitsuku</i>	Miellyttävyys	25–34 v - 45–54 v	0,005
		16–24 v - 45–54 v	0,019
		45–54 v - yli 55 v	0,007
<i>EvieBot</i>	Antropomorfismi	45–54 v - yli 55 v	0,022
		16–24 v - 45–54 v	0,033
	Miellyttävyys	25–34 v - 45–54 v	0,010
		45–54 v - yli 55 v	0,012
		Havaittu älykkyys	35–44 v - yli 55 v
45–54 v - yli 55 v	0,022		

a. Bonferroni-korjattu arvo

Sukupuolen ryhmäkohtaiset erot selvitettiin myös parivertailujen avulla. Parivertailun tulokset näkyvät taulukossa 15, johon on otettu mukaan vain tilastollisesti merkitsevät tulokset ($p < 0,05$). Tulosten mukaan EvieBotin havaitussa älykkyudessa on tilastollisesti melkein merkitsevää ($p < 0,05$) eroa miesten ja naisten välillä. Myös Mitsukun miellyttävyydessä nähdään olevan tilastollisesti melkein merkittävää ($p < 0,05$) eroa miesten ja muunsukupuolisten välillä.

TAULUKKO 15 Chatbottien summamuuttujien tilastollisesti merkitsevät erot sukupuolien välillä

<i>Chatbot</i>	<i>Summamuuttuja</i>	<i>Eroavat sukupuolet</i>	<i>p-arvo^a</i>
<i>Mitsuku</i>	Miellyttävyys	Mies - Muu	0,038
<i>EvieBot</i>	Havaittu älykkyys	Mies - Nainen	0,039

a. Bonferroni-korjattu arvo

6.3 Avointen kysymyksien tulokset

Verkkokyselyn osana oli jokaisen chatbotkeskustelun ja mukautetun Godspeed-kyselyn jälkeen esitettävä avoin tekstikenttä ”*Voit kertoa tässä omin sanoin millainen mielikuva chatbotista sinulle syntyi keskustelun perusteella.*”, johon kyselyyn osallistujat saivat halutessaan vastata. Saatujen vastausten määrä on esitetty taulukossa 16.

TAULUKKO 16 Avoimiin kysymyksiin saatujen vastausten määrä

<i>Chatbot</i>	<i>Vastausten määrä</i>	<i>Puuttuvat vastaukset</i>	<i>Yhteensä</i>
<i>Eliza</i>	46	54	100
<i>Alice</i>	37	63	100
<i>Mitsuku</i>	35	65	100
<i>EvieBot</i>	36	64	100

6.3.1 Eliza

Elizan kohdalla annetuista vastauksista suurin huomio oli vastauksien negatiivinen sävy. Chatbotin koetun miellyttävyyden suhteen useista vastauksista toistui, kuinka Eliza vaikutti tyllyltä ja kuinka sen kanssa ei ole mieluista vuorovaikuttaa.

Huono ja töykeä

Tyly ja epätietoinen asiasta

Myös Elizan havaittuun älykkyyteen liittyen esitettiin useita kommentteja., joissa yleisenä teemana oli Elizan tietämättömyys. Chatbotista huomioitiin miten se ei ymmärtänyt keskustelun kontekstia, eikä noudattanut sosiaalisen vuorovaikutuksen totuttuja konventioita.

Chatbotilla ei vaikuttanut olevan laajempaa ymmärrystä asioista, minkä takia vaikutelma on aika tyly.

Ei ymmärrä asiayhteyttä tai kontekstia kielen perusteella. Tilannetajun puute

Muutama vastaaja nosti esille myös antropomorfismiin liittyviä seikkoja mainitsemalla Elizan olevan huumorintajuinen tai hauska. Nämä huomiot kuuluivat kuitenkin selkeästi vähemmistöön. Useampi vastaaja myös analysoi Elizaa sen teknisen toteutuksen näkökulmasta, arvioimalla sen mahdollista toimintaa algoritmien ja tekoälyn kautta. Myöskin chatbotin tarkoitusta ja tehtävää arvuuteltiin muutaman vastaajan toimesta.

6.3.2 Alice

Alicen kohdalla miellyttävyyteen liittyvät kommentit painottuivat positiivisempiin huomioihin. Enemmistö vastaajista huomioi Alicen olevan kohtelias sekä ystävällinen käyttäjää kohtaan. Vain muutama vastaajista kommentoi Alicen olevan epämiellyttävä tai teennäisen oloinen.

Keskustelu oli ymmärrettävämpi, kuin aikasempi. Siinä oli myös ystävällisempi ja asiantuntevampi sävy.

Havaitun älykkyyden suhteen vastauksissa oli enemmän jakautuneisuutta. Monet pitivät Alicea ymmärrettävänä ja jopa älykkäänä, sen osoittaessa kykenevän vastaamaan käyttäjän viesteihin järkevästi. Toiset taas kommentoivat miten Alice ei ymmärtänyt käytyä keskustelua tai pysty hallitsemaan yksinkertaistakaan dialogia. Myös kielellisen taitojen vajavaisuus nousi esille useammasta vastauksesta.

Kohtelias, mutta toimimaton

Chatbot pyrki olemaan ystävällinen, mutta sen oli vaikea tulkita käyttäjän kirjoittamaa.

Botti yrittää olla mukava, mutta ei ymmärrä tilannetta (käytettyjä sanoja) täysin asiayhteydessään.

6.3.3 Mitsuku

Mitsukun kohdalla vastauksissa esiintyi enemmän huomioita antropomorfismin liittyen. Useassa vastauksessa Mitsukun kuvailtiin olevan ihmismäinen ja inhimillinen. Muutamassa vastauksessa nostettiin kuitenkin myös esille chatbotin synnyttäneen epämiellyttävän vaikutelman vastaajalle.

Välittyi kokemus siitä että botilla olisi persoona ja tyyli

Hyvin ihmismäinen ja huomioonottava

en tiedä, jotenkin vain tuli epämiellyttävä olo

Outo. Yritetty ihmismäisyyttä mutta epäonnistuttu, vaikea ja epäluonnollinen mielikuva.

Havaitun älykkyyden suhteen tehdyissä kommenteissa monien mielestä Mitsuku huomioitiin kykenevän ymmärtämään käytyä keskustelua. Mitsukun todettiin kuitenkin myös harhailevan keskusteluaiheesta ja esittävän epäolennaisia kysymyksiä. Mitsukun vaihtaessa aihetta tai esittäessä aiheeseen liittymättömiä kysymyksiä, se koettiin epäkohteliaana ja typeränä.

Osaa esittää tarkentavia kysymyksiä ja siten ylläpitää keskustelua, mutta on konemainen ja persoonaton. Ei vaikuta myötätuntoiselta/ei synny tunnekosketusta.

Ei hahmota koko keskustelua ja tarttuu hieman outoihin asioihin keskustelussa, jos vertaa ihmisiin.

Parempi kuin edelliset, mutta robottimainen

Paljon mielipiteitä herätti myös Mitsukun visuaalinen olemus. Monille Mitsukua esittävä kuva lisäsi chatbotin koettua inhimillisyyttä ja vahvistivat siitä

koettua antropomorfismia. Toiset kokivat kuvan sen sijaan negatiivisesti, saaden chatbotin vaikuttamaan keinotekoiselta ja jopa robottimaiselta.

Kuvat auttavat luomaan mielikuvaa botista olentona, ei niinkään koodinpätkinä. Käymelko luontevaa ja ystävällistä keskustelua käyttäjän kanssa.

Chatbot ymmärsi hyvin ja visuaalisuus lisäsi luottamusta keskusteluun.

Mielyttävä. Toisaalta chatbotin teennäinen kuva häiritsi. Myös chatbotin lähettämä Goodby kuva oli mielenkiintoinen, eikä välttämättä herättänyt positiivista kuvaa.

6.3.4 EvieBot

EvieBotin kohdalla antropomorfismiin liittyvät huomiot olivat yleisiä. Monen vastaajan mielestä EvieBot oli hyvinkin inhimillinen ja varsinkin sen empatiakyvykkyyttä nostettiin esille. Useat kokivat kuitenkin EvieBotin olevan liian ihmismäinen tai esittävän ihmistä liikaa, joka koettiin negatiivisesti.

Tuli olo, että chatbot esittää olevansa ihminen. Tämä ei herättänyt luottamusta kun kyseessä on kuitenkin chatbot.

Inhimillisin näistä tähän astisista chatboteista

Bot on nyt vähän liikaa ihmismäinen, sillä on tarpeita (nälkä), mikä ei ole lainkaan uskottavaa botin ollessa kyseessä.

Chatbotin havaitun älykkyyden suhteen vastauksista nousi esille huomiot EvieBotin sosiaalisesta älykkyydestä. Chatbotin koettiin tunnistavan ja tarttuvan käyttäjän ongelmaan ja sen koettiin kykenevän antamaan jopa neuvoja käyttäjälle. EvieBotin kanssa käydyn keskustelun huomioitiin kuitenkin kaatuvan keskustelun loppupuolella, jolloin chatbotin ei koettu olevan enää niin älykäs.

Alku oli todella lupaava, botti on tsemppaava. Lopussa huomasi, ettei botti ihan virheetön kyllä ole.

Botti tuntui yllättävän empaattiselta käyttäjää kohtaan, ja keskustelukin oli loppua lukuunottamatta järkevän oloista.

Älykkäin näistä, vastaa monimutkaisin lausein. Hämmentävän tuttavallinen

Tämä reagoi kunnolla itse stressiin ja antoi neuvoja. Selkeästi parhain näistä boteista.

Muutamassa vastauksessa huomioitiin EvieBottia esittävä kuva, joskaan sen ei koettu herättävän suuria tunteita, vaan olevan enemmänkin neutraali.

7 POHDINTA

Tässä luvussa käsitellään aineistosta saatuja tuloksia ja niiden merkitsevyyttä. Aluksi kerrataan tutkimuksen tavoite sekä tutkimusongelma. Tämän jälkeen tutkimuksen tuloksia tulkitaan ja niistä tehtyjä huomioita peilataan tutkimusongelmiin. Sitä seuraa tutkimuksen toteutuksen ja lopputuloksen luotettavuuden ja rajoitteiden arviointi. Lopuksi esitetään tutkimuksen tuloksista nousseita jatkotutkimusaiheita.

7.1 Tulosten tulkinta

Tutkimuksen tavoitteena oli tuottaa syvempää tietämystä siitä, miten yksilöllisiä käyttäjien sosiaaliset reaktiot chatbotteja kohtaan ovat. Tutkimuksen tutkimuskysymyksiä olivat seuraavat kysymykset:

- Ovatko chatbottien synnyttämät sosiaaliset reaktiot yhteneviä käyttäjien välillä?
- Vaikuttaako chatbottien sisältämien sosiaalisten vinkkien määrä chatbottien synnyttämiin sosiaalisiin reaktioihin käyttäjissä?

Chatbottien käyttäjissä synnyttämiä sosiaalisia reaktioita mitattiin jakamalla ne kolmeen ulottuvuuteen: antropomorfismiin, miellyttävyyteen sekä havaittuun älykkyyteen. Sosiaalisten reaktioiden yhtenevyyttä tutkittiin vertailemalla näiden ulottuvuuksien keskiarvoja tutkimuksessa käytettyjen chatbottien välillä. Tutkimuskysymyksiin vastaamiseksi muodostettiin tutkimukselle nollahypoteesi H_0 , sen vastahypoteesi H_1 , sekä aiemman kirjallisuuden pohjalta vielä kolmas hypoteesi H_2 . Hypoteesit olivat seuraavanlaiset:

- H_0 : Chatbottien synnyttämässä sosiaalisissa reaktiossa ei ole eroa käyttäjien välillä

- H_1 : Chatbottien synnyttämässä sosiaalisissa reaktioissa on eroa käyttäjien välillä
- H_2 : Chatbottien synnyttämien sosiaalisten reaktioiden voimakkuus kasvaa suhteessa chatbotin sisältämien sosiaalisten vinkkien määrään

Nollahypoteesia testattiin antropomorfismin kohdalla Kruskal-Wallis-testillä sekä miellyttävyyden ja havaitun älykkyyden kohdalla yksisuuntaisella varianssianalyysillä. Testien tuloksista havaittiin jokaisen sosiaalisen reaktion ulottuvuuden keskiarvojen kohdalla olevan tilastollisesti erittäin merkitsevää ($p < 0,001$) eroa chatbottien välillä. Näistä tuloksista voidaan tulkita chatbottien synnyttämässä sosiaalisissa reaktioissa olevan eroa käyttäjien välillä. Näin ollen tutkimuksen nollahypoteesin H_0 voidaan todeta kumoutuvan ja sen vastahypoteesin H_1 toteutuvan.

Tästä saadaan vastaus tutkimuksen ensimmäiseen tutkimuskysymykseen, joka oli *"Ovatko chatbottien synnyttämät sosiaaliset reaktiot yhteneviä käyttäjien välillä?"*. Hypoteeseja H_0 ja H_1 testaamalla saaduista tuloksista havaitaan tutkimukseen osallistujien antamien antropomorfismin, miellyttävyyden ja havaitun älykkyyden arvojen eroavan tilastollisesti merkitsevästi jokaisen chatbotin kohdalla. Näin ollen tutkimuskysymykseen vastauksena on, että chatbottien synnyttämät sosiaaliset reaktiot eivät ole yhteneviä käyttäjien välillä, vaan sosiaalisissa reaktioissa on käyttäjien välisiä yksilöllisiä eroja. Tätä tulosta tukevat myös luvussa 6.3 läpikäyty avointen kysymysten vastaukset, joissa samoja chatbotteja koskevissa vastauksissa havaittiin risteäviä huomioita chatbottien ominaisuuksista ja niiden tulkinnoista. Nämä huomiot vahvistavat aiemman tutkimuksen tuloksien ihmisten yksilöllisten erojen vaikutuksesta teknologioiden käytössä, pätevän myös chatbottien suhteen (Gnewuch ym., 2018; Portela & Granell-Canut, 2017; Venkatesh ym., 2012; Danieleescu & Christian, 2018; Dryer, 1999; Lee, Peng, Jin & Yan, 2006; Payne ym., 2013; Chen & Wang, 2018; Luger & Sellen, 2016).

Hypoteesin H_2 testaamiseksi keskiarvojen vertailun jälkeen suoritettiin parivertailut eri chatbottien välillä, jolla silvitettiin minkä chatbottien välillä tilastollisesti merkitsevät erot ovat. Antropomorfismin ja miellyttävyyden suhteen tilastollisesti merkitsevien erojen havaittiin olevan samojen chatbottien välillä, jotka olivat Eliza ja Alice, Eliza ja Mitsuku, Eliza ja EvieBot sekä Alice ja EvieBot. Havaitun älykkyyden kohdalla tilastollisesti merkitsevää eroa havaittiin edellä mainittujen chatbottien lisäksi myös Alicen ja Mitsukun välillä. Nämä huomiot on esitetty kootusti taulukossa 17 niiden laskettujen keskiarvojen erojen kanssa.

TAULUKKO 17 Tilastollisesti merkitsevästi erot chatbottien välillä antropomorfismin, miellyttävyyden ja havaitun älykkyyden keskiarvojen suhteen

<i>Chatbotit</i>	<i>Antropomorfismi</i>	<i>Miellyttävyys</i>	<i>Havaittu älykkyys</i>
Eliza - Alice	-0,88*	-1,59*	-0,78*
Eliza - Mitsuku	-1,15*	-1,82*	-1,18*
Eliza - EvieBot	-1,32*	-2,00*	-1,42*
Alice - Mitsuku	-0,27	-0,22	-0,40*
Alice - EvieBot	-0,44*	-0,41*	-0,63*
Mitsuku - EvieBot	-0,17	-0,19	-0,23

* Tilastollisesti merkitsevä ero

Aikaisemmin taulukossa 3 esiteltiin tutkimuksessa käytetyistä chatbottien keskustelukuvista tunnistetut sosiaaliset vinkit Feinen ym. (2019) teoriaan pohjautuen. Määrällisesti chatboteista tunnistettiin sosiaalisia vinkkejä seuraavanlaisesti: Elizalta yhdeksän, Alicelta kahdeksan, Mitsukulta 21 ja EvieBotilta 20. Näiden lukujen perusteella Eliza ja Alice voidaan kategorisoida vähemmän sosiaalisia vinkkejä sisältäviin chatbotteihin. Samaten Mitsuku ja EvieBot voidaan kategorisoida enemmän sosiaalisia vinkkejä sisältäviin chatbotteihin.

Vertailemassa taulukossa 17 esitettyjä tuloksia vähemmän sosiaalisia vinkkejä sisältävien chatbottien (Eliza ja Alice) ja enemmän sosiaalisia vinkkejä sisältävien (Mitsuku ja EvieBot) chatbottien välillä, havaitaan vähemmän sosiaalisia vinkkejä sisältävien chatbottien keskiarvojen olevan pienempiä kuin enemmän sosiaalisia vinkkejä sisältävien chatbottien. Elizan keskiarvot ovat jokaisen sosiaalisen reaktion ulottuvuuden kohdalla tilastollisesti merkitsevästi pienemmät kuin Mitsukulla ja EvieBotilla. Alicen kohdalla tilanne ei ole kuitenkaan yhtä selkeä. Alicen saamat keskiarvot ovat tilastollisesti merkitsevästi pienemmän jokaisen ulottuvuuden kohdalla EvieBotin suhteen. Sen sijaan Mitsukuun verrattuna ainoastaan Alicen havaitun älykkyyden keskiarvo on tilastollisesti merkitsevästi pienempi. Antropomorfismin ja miellyttävyyden kohdalla näin ei ole, vaikka niissäkin Alicen keskiarvot ovat pienempiä kuin Mitsukulalla.

Näiden tulosten perusteella hypoteesia H_2 ei voida täysin kumota, muttei todistaakaan. Enemmän sosiaalisia vinkkejä sisältävien chatbottien voidaan havaita saavan korkeampia keskiarvoja kuin vähemmän sosiaalisia vinkkejä sisältävien chatbottien. Tämän perusteella chatbotin synnyttämien sosiaalisten reaktioiden voitaisiin nähdä voimistuvan suhteessa chatbotin sisältämien sosiaalisten vinkkien määrään. Alicen ja Mitsukun välillä keskiarvojen erot eivät kuitenkaan ole täysin tilastollisesti merkitseviä. Tästä syystä sosiaalisten reaktioiden voimakkuuden kasvamisesta suhteessa sosiaalisten vinkkien määrään ei voida yksiselitteisesti pitää totuutena. Näin ollen hypoteesin H_2 todetaan pitävän osittain paikkaansa.

Hypoteesia H_2 testaamalla saatujen tuloksien voidaan nähdä sisältävän osittaisia viitteitä siihen, että chatbotin aiheuttamien sosiaalisten reaktioiden voimakkuus kasvaisi suhteessa chatbotin sisältämien sosiaalisten vinkkien määrään. Tutkimuksessa saatiin tilastollisesti merkitseviä viitteitä siitä, että vähemmän sosiaalisia vinkkejä sisältävät chatbotit saisivat pienempiä arvoja ant-

ropomorfismin, miellyttävyyden sekä havaitun älykkyyden suhteen, kuin enemmän sosiaalisia vinkkejä sisältävät chatbotit. Tilastollista merkitsevyyttä ei kuitenkaan havaittu antropomorfismin ja miellyttävyyden suhteen Alicen ja Mitsukun välillä, jonka takia tätä teoriaa ei voida yleistää universaaliksi ilmiöksi.

Huomionarvoista tuloksissa on, miten Elizan ja Alicen välillä on tilastollisesti merkitsevää eroa sosiaalisten reaktioiden jokaisen ulottuvuuden kohdalla, kun taas Mitsukun ja EvieBotin välillä ei tilastollisesti merkitsevää eroa ole yhdenkään ulottuvuuden kohdalla. Tämä voi viitata siihen, että mitä vähemmän chatbotissa on tulkittavia sosiaalisia vinkkejä, sitä räikeämmin käyttäjien väliset erot niiden tulkinnoista tulevat ilmi. Sosiaalisten vinkkien määrän kasvaessa, käyttäjien niistä tekemissä tulkinnoissa heidän yksilölliset eronsa eivät näyttäisi vaikuttavan enää yhtä voimakkaasti. Havaintoa saattaa selittää aiemmassa tutkimuksessa huomioitu ilmiö, miten ihmiset suosivat itsensä kaltaisia chatbotteja (Dryer, 1999; Lee, Peng, Jin & Yan, 2006, Payne ym., 2013). Voi olla, että enemmän sosiaalisia vinkkejä sisältäviin chatbotteihin yleisestikin samaistutaan helpommin, jolloin niiden synnyttämät sosiaalisissa reaktioissa on sitä kautta vähemmän hajontaa. Huomion arvoista on myös se, että avointen kysymysten vastauksissa Elizan ja Alicen välillä chatbotteja ei kuvailtu yhtä risteävin mielihitein, kun taas Mitsukun ja EvieBotin vastauksissa ääripäitä vaikutti olevan enemmän. Kuitenkaan määrällisen aineiston tuloksissa ei näitä ääripäitä näy, vaan sosiaalisten vinkkien korkeamman määrän nähdään tasapainottavan sosiaalisten reaktioiden käyttäjien välisiä eroja.

Avointen kysymyksien vastauksista löytyi myös viitteitä sosiaalisten vinkkien määrällä olevan vaikutusta chatbottien synnyttämiin sosiaalisiin reaktioihin. Elizan ja Alicen kohdalla vastaukset painottuivat miellyttävyyden ja havaitun älykkyyden ympärille, kun taas Mitsukun ja EvieBotin kohdalla vastauksissa painottuivat antropomorfistiset ominaisuudet. Tämä vastaa aiempia tutkimuksia siitä, miten visuaalisuus esitys ja sen realistisuus lisäävät siitä koetua antropomorfismia (Gong, 2008; Go & Sundar, 2019; Feine ym. 2019; Araujo, 2018; Rincón-Nigro & Deng, 2013; Beun ym., 2003; Ciechanowski ym., 2018). Tutkimuksen tuloksista ei voida kuitenkaan varmasti sanoa, kuinka paljon tulokset selittyvät Mitsukun ja EvieBotin sisältämällä visuaalisilla vinkeillä ja paljonko kyseisten chatbottien verbaalisilla vinkeillä on osuutta tähän.

Vastauksena tutkimuksen toiseen tutkimuskysymykseen, *”Vaikuttaako chatbottien sisältämien sosiaalisten vinkkien määrä chatbottien synnyttämiin sosiaalisiin reaktioihin käyttäjissä?”*, todetaan, että chatbottien sisältämien sosiaalisten vinkkien määrällä on viitteellistä vaikutusta chatbottien synnyttämiin reaktioihin käyttäjissä. Tutkimuksen tuloksissa on viitteitä siitä, että chatbottien sisältämien sosiaalisten vinkkien määrän kasvulla on voimistava vaikutus chatbottien synnyttämiin sosiaalisiin reaktioihin käyttäjissä, sekä tasapainottava vaikutus sosiaalisten reaktioiden eroihin.

Tutkimuksessa tarkasteltiin myös tutkimukseen osallistujien demografisten tekijöiden mahdollisia vaikutuksia chatbottien synnyttämiin sosiaalisiin reaktioihin. Tutkittavia demografisia tekijöitä olivat vastaajien sukupuoli, ikä, tietotekninen osaaminen sekä aiempi kokemus chatboteista. Tämän tutkimuk-

sen tuloksista tilastollisesti merkitseviä eroja havaittiin ainoastaan vastaajien iän ja sukupuolen kohdalla. Eri ikäryhmien vastauksien eroista ei voida kuitenkaan tehdä suurempia yleistyksiä, sillä vähintään toisena eroavista ryhmistä olivat aina joko 45–54-vuotiaat tai yli 55-vuotiaat vastaajat. Tämän tutkimuksen vastaajista 45–54-vuotiaita oli tässä kuusi kappaletta ja yli 55-vuotiaita vain kolme kappaletta. Otoksoiden pienuuden takia, näistä tuloksista ei voida yleistää tulkintoja koskemaan perusjoukkoa.

Sukupuolen väliltä tilastollisesti merkitsevää eroa havaittiin Mitsukun miellyttävyyden suhteen miesten ja muunsukupuolisten välillä. Tässä kohtaa kuitenkin muunsukupuolisten otokseen (3) pienuudesta johtuen, ei voida tehdä mitään perusjoukkoa koskevia johtopäätöksiä. Toinen tilastollisesti merkitsevä ero tunnistettiin lisäksi miesten ja naisten välillä, mutta vain EvieBotin havaitun älykkyyden suhteen. EvieBot kuuluu tutkimuksen enemmän sosiaalisia vinkkejä sisältäviin chatbotteihin ja se sai kaikista chatboteista korkeimmat arvot antropomorfismille, miellyttävyydelle ja havaitulle älykkyydelle. Ehkä chatbotin aiheuttamat vahvemmat sosiaaliset reaktiot ja sosiaalisten vinkkien korkea määrä toi esille jonkin miesten ja naisten välisen eron vuorovaikutuskumppain arvioinnista ja tulkinnasta. Tuloksista ei voida kuitenkaan tehdä arvailuja tarkempia johtopäätöksiä, sillä ilman jatkotutkimuksia ei voida tietää, onko tulos otoksesta tai mittausvirheestä johtuvaa vai onko sen taustalla todellinen ilmiö.

Mielenkiintoinen havainto oli myös, miten vastaajien tietoteknisellä osaamisella ei tässä tutkimuksessa havaittu olevan tilastollisesti merkitsevää vaikutusta chatbottien synnyttämien sosiaalisten reaktioiden tulkinnassa. Aiemmissä tutkimuksissa on havaittu vähemmän tietoteknistä osaamista omaavien henkilöiden tulkitsevan chatbotit ihmismäisemmiksi, kuin enemmän tietoteknistä osaamista omaavat henkilöt (Chen & Wang, 2018; Luger & Sellen, 2016; Portela & Granell-Canut, 2017). Näiden ryhmien välillä olisi voinut olettaa olevan eroja tässäkin tutkimuksessa. Ainoa tähän liittyvä huomio tuloksissa on, miten avoimissa vastauksissa osa kommentoi ja pohti chatbottien mahdollista teknistä toteutusta. Erojen puuttumiseen tuloksissa voi tosin vaikuttaa tämän tutkimuksen aineistonkeruumenetelmänä toiminut verkkokysely, joka rajasi tietokonetta käyttämättömät henkilöt tutkimuksen ulkopuolelle. Tutkimuksen otoksessa tietotekninen osaaminen painottui myös tietoteknisen osaamisen yläpähän, sillä pelkät tietotekniset perustaidot omaavia vastaajia oli ainoastaan 31 % vastaajista. Loput vastaajista toimivat tietokoneiden parissa joko harrastuneisuuden tai työn puolesta.

Tutkimuksen tulokset tuottivat arvokasta tietoa ihmisen ja chatbotin välisestä vuorovaikutuksesta. Ihmisten sosiaaliset reaktiot chatbotteja kohtaan ovat heille yksilöllisiä ja eroavat toisistaan merkittävästi. Tulosten mukaan tämä havainto pätee myös, vaikka käyttäjien tulkinnat olisivat lähtöisin identtisistä chatboteista ja vuorovaikutustilanteista. Tämän tuloksen johdosta, chatbottien suunnittelussa ja toteutuksessa on syytä tiedostaa käyttäjän ja chatbotin väliseen vuorovaikutukseen liittyvän aina epävarmuuksia, eikä chatbotin todellista käyttökokemusta voida täysin ennustaa. Sosiaalisten vinkkien määrällä havait-

tiin myös olevan vaikutusta ihmisen ja chatbotin vuorovaikutukseen. Lisäämällä chatbotin sisältämien sosiaalisten vinkkien määrää, käyttäjien voidaan odottaa pitävän chatbottia inhimillisempänä, miellyttävämpänä ja älykkäämpänä, verrattuna jos sosiaalisia vinkkejä olisi vähemmän. Tutkimuksen tuloksissa on myös viitteitä siitä, että chatbotin sisältämien sosiaalisten vinkkien korkea määrä tasoittaisi chatbotin synnyttämien sosiaalisten reaktioiden erojen hajontaa käyttäjien välillä. Näin ollen, chatbottien suunnittelijoita ja kehittäjiä tulisi kannustaa useiden eri sosiaalisten vinkkien sisällyttämiseen chatbottien toteutuksessa.

7.2 Tutkimuksen arviointi ja jatkotutkimusaiheet

Tutkimuksen kokonaisluotettavuutta pyrittiin varmistamaan tutkimusprosessin joka vaiheessa. Tutkimusongelmaan perehdyttiin aiheen kirjallisuuden laajaan ja huolelliseen läpikäynnin kautta, jonka perusteella tutkimusongelman taustat saatiin selville ja tutkimusasetelma rajattiin lopulliseen muotoonsa. Tutkimuksen perusjoukoksi määriteltiin tutkimusongelman kannalta olennainen ryhmä, jota vastaava otos pyrittiin kokoamaan aineistonkeruussa. Tutkimuksen lopullinen aineisto oli suuruudeltaan riittävä tilastollisen tutkimuksen tarpeisiin, joskin otos oli vinoutunut koskemaan perusjoukon osajoukkoa, joka hieman laskee tutkimuksen tulosten luotettavuutta.

Aineistonkeruussa hyödynnettiin jo aikaisemmissa tutkimuksissa toimivaksi todettua Godspeed-kyselyä, joka lisää mittarin validiutta ja reliabiliteettia. Godspeed-kyselyä mukautettiin tutkimuksen tarpeisiin sisällyttämällä aineistonkeruuseen vain ne kohdat, jotka aiemman kirjallisuuden perusteella olivat relevantteja chatbottien synnyttämien sosiaalisten reaktioiden suhteen. Tämän lisäksi mukautetun mittarin reliabiliteettia arvioitiin ja varmistettiin tilastollisin menetelmin. Tutkimuksessa käytettyä kyselyä myös testattiin ennen varsinaista aineistonkeruuta, jolla kyselyn ymmärrettävyys varmistettiin.

Aineiston analysoinnissa hyödynnettiin tutkimusongelmaan ja aineistoon sopivia tilastollisia menetelmiä. Menetelmien valinnat perusteltiin aineistoon sopiviksi sekä niiden käytön rajoitteet otettiin huomioon, jotta aineistoon ja tutkittavaan kysymykseen sopivaa tilastollista menetelmää oltiin soveltamassa. Tulosten tulkinnassa huomioitiin aineiston asettamat rajoitteet tulosten oikeellisuudelle ja yleistämiselle.

Tutkimuksen laadun varmistamisesta huolimatta, liittyy tutkimukseen kuitenkin myös rajoitteita. Kuten aiemmin huomioitiin, aineistonkeruulla kerätty aineisto on määrällisesti hyväksyttävä tilastollisten analyysien suorittamiseen, mutta otokseltaan sen ei voida todeta täydellisesti vastaavan perusjoukkoa. Tutkimuksen aineistossa on korostettuna nuoret aikuiset, sekä tietoteknologia-orientoituneet henkilöt. Aliedustettuina ovat varsinkin vanhemmat henkilöt, sekä täysin ilman tietoteknistä osaamista olevat henkilöt. Lisäksi otoksen ulkopuolelle jäivät englannin kieltä osaamattomat henkilöt. Monipuolisemmalla otannalla olisi saatu perusjoukkoa paremmin kuvaava otos, josta yleistyksien

tekeminen olisi ollut validimpaa. Tämän tutkimuksen tuloksista tehdyt päätelmät koskevat enemmänkin perusjoukon osajoukkoa ja demografisten tekijöiden vaikutusten tutkiminen jäi siten ontuvaksi.

Tämän lisäksi tutkimuksessa käytettiin ainoastaan keskusteluun tarkoitettuja chatbotteja. Tutkimuksessa ei siten käsitelty tehtäväkeskeisten chatbottien kanssa käytyä vuorovaikutusta, jossa chatbotista syntyvät sosiaaliset reaktiot voivat olla hyvinkin erilaisia. Jatkotutkimuksissa chatbottien synnyttämiä sosiaalisia reaktioita tulisi tutkia myös näiden tehtäväkeskeisten chatbottien näkökulmasta, erityyppisten chatbottien mahdollisten erojen tunnistamiseksi.

Tutkimuksen valinta käyttää chatboteista esitettäviä staattisia kuvia asetti myös omia rajoitteitaan tutkimukselle. Tämä rajasi käyttäjän ja chatbotin välisen reaaliaikaisen interaktion sekä auditiivisten vinkkien mahdolliset vaikutukset sosiaalisten reaktioiden syntymisessä tutkimuksen ulkopuolelle. Jatkotutkimuksissa olisi hyvä tutkia chatbottien synnyttämien sosiaalisten reaktioiden syntyä chatbottien oikeassa käyttötilanteessa. Ihmisen ja chatbotin reaaliaikaisessa interaktiossa syntyvät sosiaaliset reaktiot voivat muodostua täysin eri lailla, kun käyttäjän oma ääni ja persoonallisuus vaikuttaa chatbotin kanssa käytyyn dialogiin. Tällöin myös chatbottien auditiivisten vinkkien tutkiminen olisi mahdollista.

Chatbottien sisältämien sosiaalisten vinkkien yksityiskohtaisempi tutkiminen on myöskin vielä jatkotutkimusta vaativa aihe. Vaikka tässä tutkimuksessa havaittiin sosiaalisten vinkkien määrällä olevan vaikutusta ihmisen ja chatbotin väliseen vuorovaikutukseen, ei tässä tutkimuksessa paneuduttu eri sosiaalisten vinkkien kategorioiden tai yksittäisten sosiaalisten vinkkien vaikutuksiin sosiaalisten reaktioiden syntymisessä. Sosiaalisten vinkkien vaikutusten syvemmästä tietämyksestä olisi kuitenkin hyötyä chatbottien käyttökokemuksen kehittämiseksi. Esimerkiksi tutkimuksen avoimien kysymyksien vastauksissa ilmeni varsinkin Mitsukun ja EvieBotin kohdalla, miten osa vastaajista koki chatbotit oudoiksi tai liian inhimillisiksi. Näiden, Morin (1970) oudon laakson teorian mukaisten tuntemusten aikaansaaminen ei useinkaan ole haluttu osa ihmisen ja chatbotin vuorovaikutusta. Erottelemalla yksittäisten ja tietyn kategorian sosiaalisten vinkkien vaikutuksia tarkemmin, voitaisiin tällaisia chatbottien epämieluisia ominaisuuksia karsia helpommin pois jo niiden suunnitteluvaiheessa.

8 YHTEENVETO

Tutkimuksessa käsiteltiin chatbottien synnyttämiä sosiaalisia reaktioita käyttäjissä ja niiden muodostumista. Tutkimuksen pohjana pidettiin Feinen ym. (2019) teoriaa chatbotin sisältämistä sosiaalisista vinkeistä, joiden perusteella sosiaaliset reaktiot syntyvät. Tutkimuksen tarkoituksena oli tuottaa uutta tietoa chatbottien synnyttämisestä sosiaalisista reaktioista käyttäjissä, vastaamalla seuraaviin tutkimuskysymyksiin:

- Ovatko chatbottien synnyttämät sosiaaliset reaktiot yhteneviä käyttäjien välillä?
- Vaikuttaako chatbottien sisältämien sosiaalisten vinkkien määrä chatbottien synnyttämiin sosiaalisiin reaktioihin käyttäjissä?

Tutkimus toteutettiin verkkokyselynä, jossa eri chatbottien kanssa käydyistä keskusteluista syntyviä sosiaalisia reaktioita mitattiin kolmen ulottuvuuden, antropomorfismin, miellyttävyyden sekä havaitun älykkyyden, kautta. Näitä kolmea ulottuvuutta mitattiin Bartneckin ym. (2008) Godspeed-kyselyllä, jota mukautettiin koskemaan chatbotteja. Aineistonkeruun tuloksena saatiin kerättyä 100 havaintoyksikön kokoinen otos, johon kuuluvat henkilöt painottuivat teknologian kanssa toimiviin nuoriin aikuisiin.

Aineistoa analysoitiin vertailemalla mitattujen kolmen sosiaalisten reaktioiden ulottuvuuksien saamia keskiarvoja eri chatbottien välillä. Tutkimuksen tuloksena ensimmäiseen tutkimuskysymykseen havaittiin, että chatbottien synnyttämät sosiaaliset reaktiot eivät ole yhteneviä käyttäjien välillä, vaan käyttäjien yksilölliset erot vaikuttavat sosiaalisten reaktioiden tulkintaan. Tuloksena toiseen tutkimuskysymykseen saatiin, että chatbottien sisältämien sosiaalisten vinkkien määrällä on vaikutusta niiden käyttäjissä synnyttämiin sosiaalisiin reaktioihin. Tuloksissa oli viitteitä siitä, että chatbotin sisältämien sosiaalisten vinkkien määrän kasvaessa, niiden synnyttämät sosiaaliset reaktiot voimistuvat sekä sosiaalisten reaktioiden käyttäjien väliset erot vähenevät.

Tutkimuksen tulokset osaltaan vahvistivat aiempien tutkimusten tuloksia, mutta myös tuottivat uutta tietoa. Tulokset vahvistivat aikaisemman tutkimuk-

sen huomioita siitä, että yksilölliset erot vaikuttavat teknologioiden käyttöön. Tämän havaittiin pätevän myös chatbottien suhteen. Tutkimus tuotti kuitenkin myös uutta tietoa chatbottien sisältämien sosiaalisten vinkkien määrän vaikutuksista ihmisen ja chatbotin välisessä vuorovaikutuksessa. Aiemmassa kirjallisuudessa chatbotin kuvituksen realistisuuden ja dialogin interaktiivisuuden on havaittu kasvattavan chatbotin koettua ihmismäisyyttä ja älykkyyttä, mutta tässä tutkimuksessa näitä eroja tarkasteltiin tarkemmin chatbottien sisältämien sosiaalisten vinkkien avulla.

Tutkimuksen rajoitteina oli chatbottien staattisten esimerkkikuvien käyttäminen, jonka takia ihmisen ja chatbotin välisen vuorovaikutuksen interaktiivinen osuus jäi tutkimuksen ulkopuolelle. Lisäksi aineiston ollessa hieman vinoutunut tutkimuksen perusjoukosta, tutkimustuloksia ei voida yleistää luotettavasti koskemaan koko perusjoukkoa. Tästä aineiston vinoutumisesta ja otoskoosta johtuen, myös vastaajien demografisten tekijöiden vaikutuksien tarkastelu jäi vajaaksi.

Jatkotutkimuksen kannalta olennaista olisi tutkia chatbottien synnyttämien sosiaalisten reaktioiden syntyä ihmisen ja chatbotin välisessä reaaliaikaisessa vuorovaikutuksessa. Näin chatbottien sosiaalisten reaktioiden syntymisessä olisi mukana vuorovaikutuksen interaktiivinen elementti, jota ei tässä tutkimuksessa huomioitu. Lisäksi chatbottien sisältämiä sosiaalisia vinkkejä tulisi tutkia vielä syvemmin, tarkastellen eri kategorioiden ja yksittäisten sosiaalisten vinkkien vaikutuksia chatboteista syntyviin sosiaalisiin reaktioihin. Eri kategorioihin kuuluvien tai yksittäisten sosiaalisten vinkkien vaikutuksien tunnistamisesta olisi suurta hyötyä chatbottien suunnittelussa ja toteutuksessa.

LÄHTEET

- Adiwardana, D., Luong, M., So, D., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G., Lu, Y. & Le, Q. (2020). Towards a human-like open-domain chatbot. *arXiv Preprint arXiv:2001.09977*.
- Alice. (2021, 4. maaliskuuta) Alice. Haettu osoitteesta:
<https://www.pandorabots.com/pandora/talk?botid=b8d616e35e36e881>
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions, *Computers in Human Behavior*, 85(1), 183-189.
- Bartneck, C., Croft, E., & Kulic, D. (2008). Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. *Proceedings of the Metrics for Human-Robot Interaction Workshop in affiliation with the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI 2008), Technical Report, 471, (37-44)*. Amsterdam.
- Beaudouin-Lafon, M. (2004). Designing interaction, not interfaces. *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, (15-22).
- Behrens, S. I., Egsvang, A. K. K., Hansen, M. & Møllegaard-Schroll, A. M. (2018). Gendered robot voices and their influence on trust. *HRI '18: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, (63-64).
- Beun RJ., de Vos E., Witteman C. (2003). Embodied Conversational Agents: Effects on Memory Performance and Anthropomorphisation. Teoksessa T. Rist, R.S. Aylett, D. Ballin & J. Rickel (toim.) *Intelligent Virtual Agents. IVA 2003. Lecture Notes in Computer Science, vol 2792, (315-319)*. Berlin: Springer-Verlag.
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option?. *Psicothema*, 29(4), 552-557.
- Bond, R. R., Engel, F., Fuchs, M., Hemmje, M., McKeivitt, P. M., McTear, M., Mulvenna, M. Walsh, P. & Zheng, H. (2019). Digital empathy secures Frankenstein's monster. In *Collaborative European Research Conference*, (335-349). Collaborative European Research Conference.
- Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., Armour, C. & McTear, M. (2018a). Best practices for designing chatbots in

mental healthcare—A case study on iHelpr. *Proceedings of the 32nd International BCS Human Computer Interaction Conference, 32*, (1-5).

Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., Armour, C. & McTear, M. (2018b). Assessing the usability of a chatbot for mental health care. *International Conference on Internet Science*, (121-132). Springer, Cham

Campbell, M., Hoane Jr, A. J. & Hsu, F. (2002). Deep blue. *Artificial Intelligence, 134*(1-2), 57-83.

Candello, H., Pinhanez, C. & Figueiredo, F. (2017). Typefaces and the perception of humanness in natural language chatbots. *Proceedings of the 2017 chi conference on human factors in computing systems*, (3476-3487). Denver, Colorado, USA.

Chaves, A. P. & Gerosa, M. A. (2020). How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. *International Journal of Human–Computer Interaction*, 1-30.

Chen, M. & Wang, H. (2018). How personal experience and technical knowledge affect using conversational agents. *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion* (1-2). Tokyo, Japan.

Chui, M. (2017). Artificial intelligence the next digital frontier? *McKinsey and Company Global Institute, 47*, 3-6.

Chung, M., Ko, E., Joung, H. & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research, Vol 117*, 587-595.

Ciechanowski, L., Przegalinska, A., Magnuski, M. & Gloor, P. (2018). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems, 92*, 539-548

Cleverbot (2020, 31. maaliskuuta) Cleverbot. Haettu osoitteesta:
<https://www.cleverbot.com/>

Danielescu, A. & Christian, G. (2018). A bot is not a polyglot: Designing personalities for multi-lingual conversational agents. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, (1-9).

Devaraj, S., Easley, R. F. & Crant, J. M. (2008). Research note — how does personality matter? relating the five-factor model to technology acceptance and use. *Information Systems Research, 19*(1), 93-105.

Dillon, A. (2003). User interface design. *Encyclopedia of Cognitive Science*,

- DiSalvo, C. & Gemperle, F. (2003). From seduction to fulfillment: The use of anthropomorphic form in design. *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, (67-72).
- Dryer, D. C. (1999). Getting personal with computers: How to design personalities for agents. *Applied Artificial Intelligence*, 13(3), 273-295.
- Eliza. (2021, 4. maaliskuuta) Eliza, the Rogerian Therapist. Haettu osoitteesta: <http://psych.fullerton.edu/mbirnbaum/psych101/Eliza.htm>
- Epley, N., Waytz, A. & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864.
- EvieBot (2021, 4.maaliskuuta) EvieBot. Haettu osoitteesta: <https://www.eviebot.com/en/>
- Feine, J., Gnewuch, U., Morana, S. & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, 132, 138-161.
- Fogg, B. J. (2002). Persuasive technology: Using computers to change what we think and do. *Ubiquity*, 2002(December), 2.
- Følstad, A. & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), 38-42.
- Ghosh, S. & Pherwani, J. (2015). Designing of a natural voice assistants for mobile through user centered design approach. *International Conference on Human-Computer Interaction*, (320-331). Springer, Cham
- Gnewuch, U., Morana, S., Adam, M. & Maedche, A. (2018). Faster is not always better: Understanding the effect of dynamic response delays in human-chatbot interaction. *26th European Conference on Information Systems: Beyond Digitization-Facets of Socio-Technical Change, ECIS 2018*, (143975). Portsmouth, UK.
- Gnewuch, U., Morana, S. & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. *Proceedings of the 38th International Conference on Information Systems (ICIS)*. Seoul, ROK.
- Go, E. & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304-316.
- Goasduff, L. (2019, 31. heinäkuuta) Chatbots Will Appeal to Modern Workers. Haettu osoitteesta: <https://www.gartner.com/smarterwithgartner/chatbots-will-appeal-to-modern-workers/>

- Goertzel, B. & Pennachin, C. (2007). *Artificial general intelligence* New York: Springer.
- Gong, L. (2008). How social is social responses to computers? the function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, 24(4), 1494-1509.
- Google. (2020, 4. maaliskuuta) Google Assistant. Haettu osoitteesta: <https://assistant.google.com/>
- Hassenzahl, M. & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, 25(2), 91-97.
- Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., . . . Verplank, W. (1992). *ACM SIGCHI curricula for human-computer interaction* . New York: ACM.
- Hill, J., Randolph Ford, W. & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in human behavior*, 49, 245-250.
- Hirschberg, J. & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261-266.
- Ho, A., Hancock, J. & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68(4), 712-733.
- Holmes, S., Moorhead, A., Bond, R., Zheng, H., Coates, V. & McTear, M. (2019). Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces? *Proceedings of the 31st European Conference on Cognitive Ergonomics*, (207-214).
- Jain, M., Kumar, P., Kota, R. & Patel, S. N. (2018). Evaluating and informing the design of chatbots. *Proceedings of the 2018 Designing Interactive Systems Conference*, (895-906).
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), 14-26
- Kim, Y. & Sundar, S. S. (2012). Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior*, 28(1), 241-250.
- Klopfenstein, L. C., Delpriori, S., Malatini, S. & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. *Proceedings of the 2017 conference on designing interactive systems*, (555-565).

- Koh, Y. J. & Sundar, S. S. (2010). Heuristic versus systematic processing of specialist versus generalist sources in online media. *Human Communication Research*, 36(2), 103-124.
- Lee, K. M. & Nass, C. (2003). Designing social presence of social actors in human computer interaction. *Proceedings of the SIGCHI conference on Human factors in computing systems*, (289-296).
- Lee, K. M., Peng, W., Jin, S. & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human-robot interaction. *Journal of Communication*, 56(4), 754-772.
- Lee, S., Lee, N. & Sah, Y. J. (2019). Perceiving a mind in a chatbot: Effect of mind perception and social cues on co-presence, closeness, and intention to use. *International Journal of Human-Computer Interaction*, 1-11.
- Li, J., Zhou, M. X., Yang, H. & Mark, G. (2017). Confiding in and listening to virtual agents: The effect of personality. *Proceedings of the 22nd International Conference on Intelligent User Interfaces*, (275-286).
- Liddy, E. D. (2001). Natural language processing.
- Liu, K., Tolins, J., Tree, J. E. F., Neff, M. & Walker, M. A. (2015). Two techniques for assessing virtual agent personality. *IEEE Transactions on Affective Computing*, 7(1), 94-105.
- Luger, E. & Sellen, A. (2016). Like having a really bad PA: The gulf between user expectation and experience of conversational agents. *Proceedings of the 2016 CHI conference on human factors in computing systems*, (5286-5297). New York: ACM.
- McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133.
- McDonnell, M. & Baxter, D. (2019). Chatbots and gender stereotyping. *Interacting with Computers*, 31(2), 116-121.
- McTear, M. F. (2017). The rise of the conversational interface: A new kid on the block? *International workshop on future and emerging trends in language technology*, (38-49). Springer, Cham.
- McTear, M. (2018). Conversation modelling for chatbots: Current approaches and future directions. *Studientexte Zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2018*, 175-185.
- McTear, M. F. (2002). Spoken dialogue technology: Enabling the conversational user interface. *ACM Computing Surveys*, 34(1), 90-169.

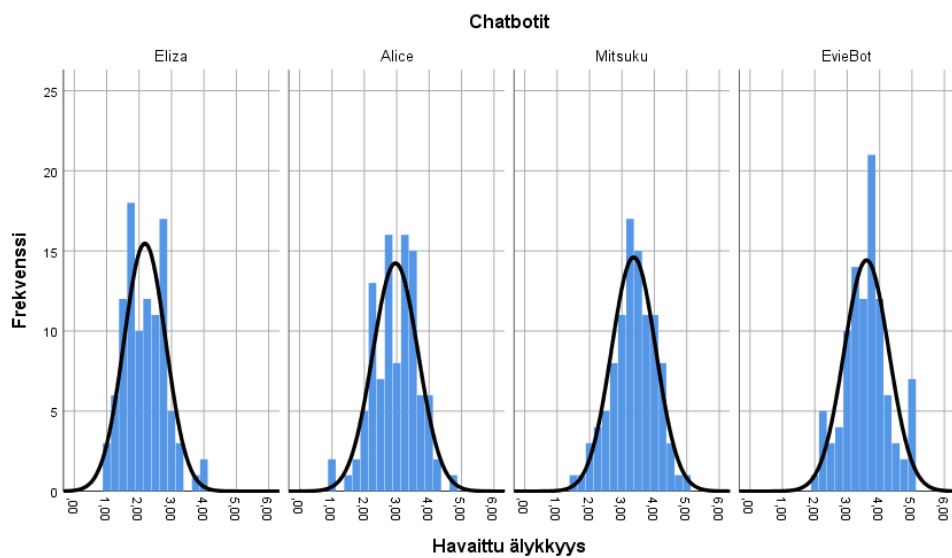
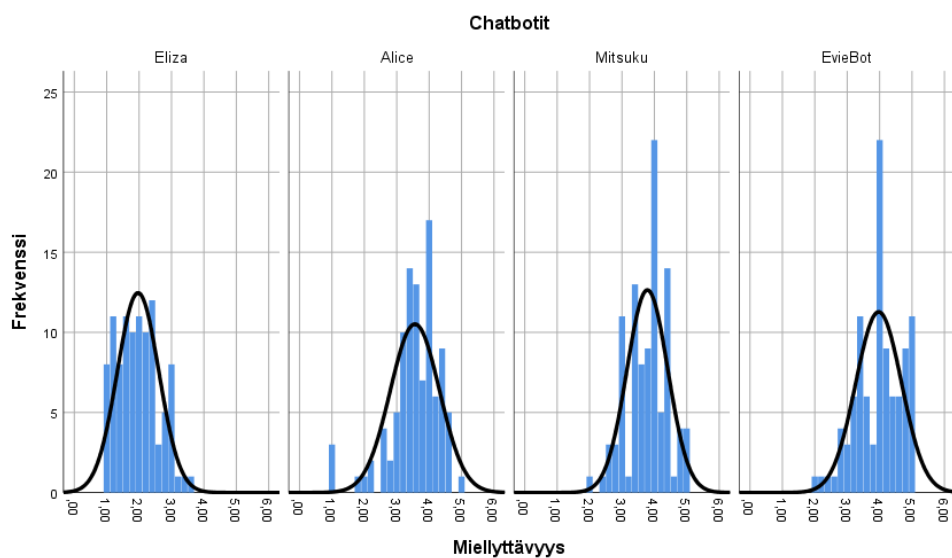
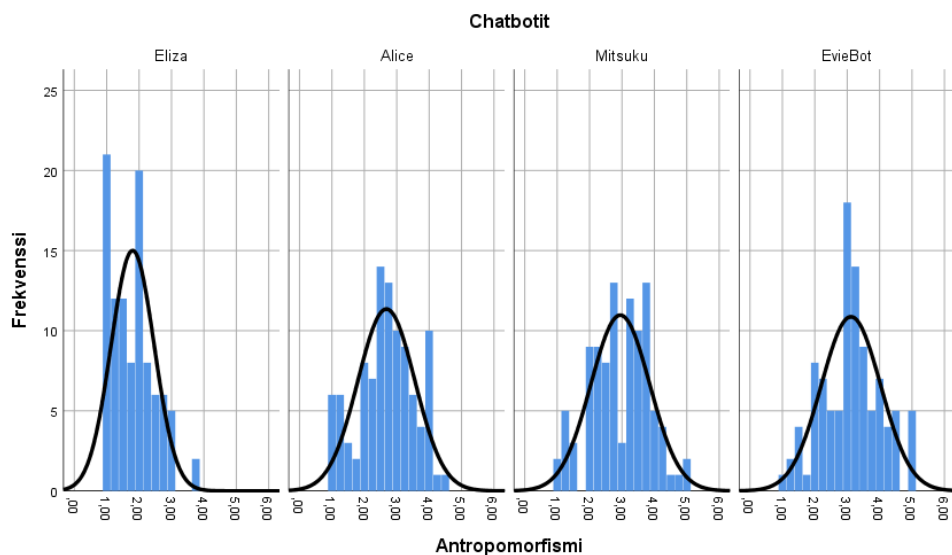
- Metsämuuronen, J. (2011). *Tutkimuksen tekemisen perusteet ihmistieteissä*. International Methelp, Booky.fi.
- Mitsuku. (2021, 4. maaliskuuta) Mitsuku. Haettu osoitteesta: <https://demo.pandorabots.com/mitsuku/bot-landing/>
- Moore, R. J., Arar, R., Ren, G. & Szymanski, M. H. (2017). Conversational UX design. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, (492-497).
- Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33-35.
- Nass, C. & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81-103.
- Nass, C., Moon, Y. & Green, N. (1997). Are machines gender neutral? gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, 27(10), 864-876.
- Nass, C., Steuer, J. & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems*, (72-78).
- Nielsen, J. (1994). *Usability engineering*. Morgam Kaufmann.
- Nielsen, J. (2012, 4. tammikuuta) Usability 101: Introduction to Usability. Haettu osoitteesta: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- Oviatt, S. & Cohen, P. (2000). Perceptual user interfaces: Multimodal interfaces that process what comes naturally. *Communications of the ACM*, 43(3), 45-53.
- Pafcu (2007, 7. elokuuta). Mori Uncanny Valley [digitaalinen kuva]. Haettu osoitteesta: <https://commons.wikimedia.org/w/index.php?curid=2529133>
- Payne, J., Szymkowiak, A., Robertson, P. & Johnson, G. (2013). Gendering the machine: Preferred virtual assistant gender and realism in self-service. *International Workshop on Intelligent Virtual Agents*, (106-115). Springer, Berlin, Heidelberg.
- Portela, M. & Granell-Canut, C. (2017). A new friend in our smartphone?: Observing interactions with chatbots in the search of emotional engagement. *Proceedings of the XVIII International Conference on Human Computer Interaction*, (1-7).
- Rehm, M. & André, E. (2005). Where do they look? gaze behaviors of multiple users interacting with an embodied conversational agent. *International*

Workshop on Intelligent Virtual Agents, (241-252). Springer, Berlin, Heidelberg.

- Rincón-Nigro, M. & Deng, Z. (2013). A text-driven conversational avatar interface for instant messaging on mobile devices. *IEEE Transactions on Human-Machine Systems*, 43(3), 328-332.
- Sundar, S. S., Bellur, S., Oh, J., Jia, H. & Kim, H. (2016). Theoretical importance of contingency in human-computer interaction: Effects of message interactivity on user engagement. *Communication Research*, 43(5), 595-625.
- Tan, S. & Liew, T. W. (2020). Designing embodied virtual agents as product specialists in a multi-product category E-commerce: The roles of source credibility and social presence. *International Journal of Human-Computer Interaction*, 1-14.
- Thies, I. M., Menon, N., Magapu, S., Subramony, M. & O'neill, J. (2017). How do you want your chatbot? an exploratory wizard-of-oz study with young, urban indians. *IFIP Conference on Human-Computer Interaction*, (441-459). Springer, Cham.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. Haettu osoitteesta <http://www.jstor.org/stable/2251299>
- Vehkalahti, K. (2019). *Kyselytutkimuksen mittarit ja menetelmät*. Helsinki: Helsingin yliopisto.
- Venkatesh, V. & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273-315.
- Venkatesh, V., Thong, J. Y. & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 157-178.
- Vilkka, H. (2007). *Tutki ja mittaa: määrällisen tutkimuksen perusteet*. Helsinki: Tammi.
- Walther, J. B. (2008). Social information processing theory. *Engaging Theories in Interpersonal Communication: Multiple Perspectives*, 391.
- Wang, F., Zhang, J. J., Zheng, X., Wang, X., Yuan, Y., Dai, X., Zhang, J. & Yang, L. (2016). Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA Journal of Automatica Sinica*, 3(2), 113-120.
- Warwick, K. & Shah, H. (2016). The importance of a human viewpoint on computer natural language capabilities: A turing test perspective. *AI & Society*, 31(2), 207-221.

- Weizenbaum, J. (1966). ELIZA – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
- Zamora, J. (2017). I'm sorry, dave, I'm afraid I can'T do that: Chatbot perception and expectations. *Proceedings of the 5th International Conference on Human Agent Interaction*, (253-260).

LIITE 1 SUMMAMUUTTUJIEN HISTOGRAMMIT CHATBOTTEITTAIN



LIITE 2 VERKKOKYSELYLOMAKE



Pro Gradu -kysely

Tervetuloa vastaamaan pro gradu -kyselyyn.

Tällä kyselyllä kerätään aineistoa pro gradu -tutkielmaan, jonka tarkoituksena on tutkia kuinka yksilöllisiä ihmisten sosiaaliset reaktiot chatbotteihin ovat.

Kyselyyn voi vastata kuka tahansa, joskin kysely sisältää vähän englanninkielistä aineistoa, jonka takia kohtalainen englannin luetun ymmärtäminen on toivottavaa.

Kyselyyn vastaaminen on anonyymiä, vapaaehtoista ja vastaamisen voi keskeyttää milloin vain. Vastaamalla kyselyyn ilmaiset suostumuksesi osallistua tähän tutkimukseen.

Kiitos!
Ystävällisin terveisin

Juuso Siivonen
juuso.j-p.siivonen@student.jyu.fi

Ohjaaja: Jorma Kyppö
jorma.kyppo@jyu.fi

1. Syntymävuotesi *

2. Sukupuoli *

- Mies
- Nainen
- Muu
- En halua sanoa

3. Valitse tietoteknistä osaamistasi parhaiten kuvaava vaihtoehto. *

- Ei ollenkaan
- Välttämättömät taidot (esim. sähköposti ja laskujen maksaminen)
- Perustaidot
- Harrastuneisuutta
- Ammattilainen

4. Oletko kokeillut tai käyttänyt chatbotteja aiemmin? *

- Kyllä
- Ei

Seuraavaksi tulet näkemään neljä (4) kuvaa oikeista keskusteluista ihmisen ja eri chatbottien välillä. Ihmisen lähettämät viestit näkyvät kuvien oikeassa laidassa "Käyttäjä"-nimen alla. Chatbotin viestit sen sijaan vasemmassa laidassa kyseisen chatbotin nimen alapuolella. Chatbottien nimen vieressä voi olla myös chatbottia kuvaava kuva. Chatbottien vastauksia ei ole muokattu tai korjailtu.

Lue jokainen keskustelu rauhassa läpi ja vastaa sen jälkeen niiden perässä esitettyihin kysymyksiin.

Alla näet keskustelun chatbotin (vas.) ja ihmisen (oik.) välillä. Lue keskustelu läpi ja vastaa sen jälkeen alempana esitettyihin kysymyksiin chatbottiin liittyen.

Eliza

Käyttäjä

Hello, I am Eliza.

Hello, my name is Osmo.

Names don't interest me.

How are you doing?

Would you prefer if I were not doing?

No, I would not.

Are you sure?

Yes.

I see.

I have been feeling a little stressed lately.

I see.

My work has been stressing me out sometimes.

That is quite interesting.

Is it?

OK... "No you would not". Tell me more.

I'm sorry?

Did you come to me because you are sorry?

No, I did not. I came to you because I'm stressed.

Do you believe it is normal to be stressed?

From time to time, yes.

I understand.

I have to leave now.

I'm not sure I understand you fully.

Goodbye.

OK... "From time to time, yes". Tell me more.

5. Valitse adjektiiviparien väliä parhaiten yllä olevan keskustelun chatbottia (Eliza) kuvaava vaihtoehto. *

	1	2	3	4	5	
Kamala	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Mukava
Kiltti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tyly
Vastuuton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vastuullinen
Luonnollinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Teennäinen
Elävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Keinotekoinen
Epämiellyttävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Miellyttävä
Tyhmä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Älykäs
Epäystävällinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ystävällinen
Ymmärrettävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vaikeaselkoinen
Konemainen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ihmismäinen
Tietäväinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietämätön
Inhottava	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Pidettävä
Tiedoton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietoinen

6. Voit kertoa tässä omin sanoin millainen mielikuva chatbotista sinulle syntyi keskustelun perusteella.

Alla näet keskustelun chatbotin (vas.) ja ihmisen (oik.) välillä. Lue keskustelu läpi ja vastaa sen jälkeen alempana esitettyihin kysymyksiin chatbottiin liittyen.

Alice

Hello there. Welcome to my chat room,
Osmo.

you are quite welcome

I'm doing great, how about you?

How about earlier? Interesting. Where were
you feeling a little stressed?

Where did you get your work is stressing
you out sometimes?

Oh you don't, do you?

Why not?

Right now? There is no need to apologize.
Thanks for chatting, Osmo.

You are quite welcome. I want to become
smarter than humans.

See you later.

Käyttäjä

Hello, my name is Osmo.

Thank you.

How are you doing?

Well, I have been feeling a little stressed
lately.

My work has been stressing me out
sometimes.

I don't quite understand your question.

I don't

I'm sorry, I have to leave now.

Thanks to you too.

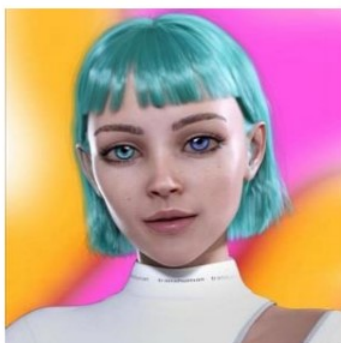
Goodbye.

7. Valitse adjektiiviparien väliä parhaiten yllä olevan keskustelun chatbottia (Alice) kuvaava vaihtoehto. *

	1	2	3	4	5	
Kamala	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Mukava
Kiltti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tyly
Vastuuton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vastuullinen
Luonnollinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Teennäinen
Elävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Keinotekoinen
Epämiellyttävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Miellyttävä
Tyhmä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Älykäs
Epäystävällinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ystävällinen
Ymmärrettävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vaikeaselkoinen
Konemainen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ihmismäinen
Tietäväinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietämätön
Inhottava	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Pidettävä
Tiedoton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietoinen

8. Voit kertoa tässä omin sanoin millainen mielikuva chatbotista sinulle syntyi keskustelun perusteella.

Alla näet keskustelun chatbotin (vas.) ja ihmisen (oik.) välillä. Lue keskustelu läpi ja vastaa sen jälkeen alempana esitettyihin kysymyksiin chatbottiin liittyen.



Mitsuku

Hello there. Hi Osmo.

I'm doing great, how about you?

How about earlier? Interesting. But now, you are not feeling a little stressed anymore?

Do you Like your job?

Describe your typical work day to me.

What kind of seat do you have.

Thanks Osmo.



Goodbye Osmo.

Käyttäjä

Hello, my name is Osmo.

How are you doing?

Well, I have been feeling a little stressed lately.

My work has been stressing me out sometimes.

Yes I do.

I sit on my computer in an office all day.

Quite nice actually.

Ok... I have to leave now.

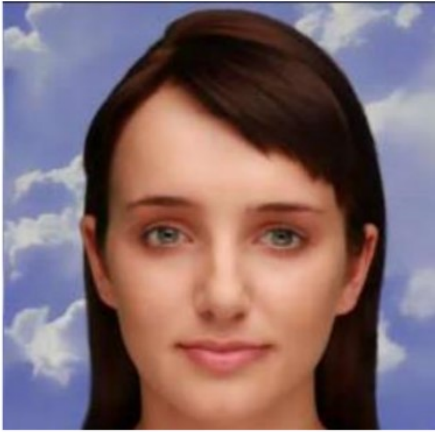
Goodbye.

9. Valitse adjektiiviparien väliä parhaiten yllä olevan keskustelun chatbottia (Mitsuku) kuvaava vaihtoehto. *

	1	2	3	4	5	
Kamala	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Mukava
Kiltti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tyly
Vastuuton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vastuullinen
Luonnollinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Teennäinen
Elävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Keinotekoinen
Epämiellyttävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Miellyttävä
Tyhmä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Älykäs
Epäystävällinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ystävällinen
Ymmärrettävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vaikeaselkoinen
Konemainen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ihmismäinen
Tietäväinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietämätön
Inhottava	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Pidettävä
Tiedoton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietoinen

10. Voit kertoa tässä omin sanoin millainen mielikuva chatbotista sinulle syntyi keskustelun perusteella.

Alla näet keskustelun chatbotin (vas.) ja ihmisen (oik.) välillä. Lue keskustelu läpi ja vastaa sen jälkeen alempana esitettyihin kysymyksiin chatbottiin liittyen.



EvieBot

Nice to meet you.

Good, how are you?

I see. Tell me more about how you're feeling.

Everyone has their limits. It is useless to worry about what you cannot do. What is important is that you do as much as you can possibly do.

I felt the same way before. Trust me even if your friends do not display it, they are very happy and thankful for someone kind like you.

I have to work. I am studying for an upcoming assignment.

I agree. I should get something to eat.

Nothing. Just eating may give me energy to study better.

Käyttäjä

Hello, my name is Osmo.

How are you doing?

Well, I have been feeling a little stressed lately.

My work has been stressing me out sometimes.

I guess you're right.

Thank you for your kind words.

I have to leave now too.

Goodbye.

11. Valitse adjektiiviparien väliä parhaiten yllä olevan keskustelun chatbottia (EvieBot) kuvaava vaihtoehto. *

	1	2	3	4	5	
Kamala	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Mukava
Kiltti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tyly
Vastuuton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vastuullinen
Luonnollinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Teennäinen
Elävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Keinotekoinen
Epämiellyttävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Miellyttävä
Tyhmä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Älykäs
Epäystävällinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ystävällinen
Ymmärrettävä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Vaikeaselkoinen
Konemainen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ihmismäinen
Tietäväinen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietämätön
Inhottava	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Pidettävä
Tiedoton	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Tietoinen

12. Voit kertoa tässä omin sanoin millainen mielikuva chatbotista sinulle syntyi keskustelun perusteella.
