

De Verschillen tussen AI en de Mens in de Klantenservice
Een onderzoek naar de ervaringen van mensen tijdens de omgang met een AI chatbot vergeleken met een menselijke chat agent.

Elena Harutunyan (2050150)

Communicatie- en Informatiewetenschappen

Specialisatie: Business Communication & Digital Media

Tilburg University

Bachelor Scriptie

Begeleider: Dr. Jan de Wit

Tweede lezer: Dr. Debby Damen

Datum: 22-06-2024

Aantal woorden: 7198

Verklaring van Technologiegebruik

Hierbij verklaar ik, Elena Harutunyan, dat deze afstudeerscriptie enkel geschreven is door mezelf, en het mijn eigen werk is.

Daarnaast verklaar ik dat in de gehele scriptie geen gebruik is gemaakt van technologische (AI) programma's. Er is enkel de AI literatuurdatabase Consensus gebruikt, wat mij heeft geholpen om meer literatuur te vinden over mijn onderzoek.

Abstract

De wereld van de klantenservice is zichtbaar aan het veranderen. Op steeds meer vlakken wordt de hulpverlening van AI chatbots ingezet, om menselijke medewerkers te ondersteunen of juist compleet te vervangen. Om de onderzoeksvraag “Hoe ervaren personen de omgang met een AI chatbot vergeleken met een menselijke chat agent in de klantenservice?” te beantwoorden, is deze literatuurstudie uitgevoerd. Overwegend hebben de resultaten van dit onderzoek aangetoond dat gebruikers een voorkeur hebben voor een menselijke chat agent, vergeleken met een AI chatbot. In het geval dat gebruikers een gesprek moeten voeren met een chatbot, hebben zij een voorkeur voor een embodied chatbot met antropomorfische eigenschappen, vergeleken met een disembodied chatbot zonder antropomorfische eigenschappen. De bevindingen van dit onderzoek kunnen de kenniskloof verkleinen met betrekking tot het begrijpen van de voorkeuren van mensen voor chatbots of menselijke chat agents. Daarnaast kunnen de onderzoeksresultaten bedrijven en beleidsmakers helpen om chatbots en menselijke agents te implementeren in de klantenservice, op basis van de voorkeuren en verwachtingen van gebruikers.

Inhoudsopgave

1. Inleiding.....	4
2. Theoretisch kader	6
2.1 Live chats en de behoeftes van mensen.....	6
2.2 Kunstmatige Intelligentie als een klantenservice-chatbot	7
2.3 De verschillen tussen AI en een menselijke chat agent.....	10
3. Methode	12
3.1 Databases en selectiecriteria.....	12
3.2 Zoektermen en overige zoekstrategieën	12
3.3 Data-analyse	13
4. Resultaten.....	14
4.1 Embodied en disembodied chatbots	14
4.2 AI chatbots en menselijke chat agents.....	18
Tabel 1	20
5. Discussie en conclusie.....	31
5.1 Discussie.....	31
5.2 Theoretische en praktische implicaties.....	33
5.3 Limitaties en suggesties voor vervolgonderzoek	33
5.4 Conclusie	34
6. Referentielijst.....	35

1. Inleiding

De hedendaagse klantenservice is anders dan vroeger. Deze heeft zich bijvoorbeeld aangepast aan de digitale wereld, met de toenemende *live chats* die een groot deel van de telefonische gesprekken hebben vervangen (Leggett, 2017). Live chats bieden klanten de mogelijkheid om via een chat contact te hebben met een *chat agent*, meestal een medewerker van de klantenservice, via de website van het bedrijf. Dit is een alternatief voor het telefonisch contact, dat tot op heden het primaire voorbeeld was van klantenservice op afstand (Leggett, 2017).

Volgens het onderzoek van Sanglé-Ferrière en Voyer (2019) zijn aan een live chat enkele voordelen verbonden, vergeleken met telefonisch contact; bijvoorbeeld een kortere responstijd, een verminderd gevoel van schaamte om vragen te stellen en gemakkelijker contact. Daarnaast zijn ook enkele mogelijke nadelen aanwezig. Het onderzoek van Helms en Mayo (2008) toonde aan dat er in sommige situaties een lange wachttijd was, de live chats niet in de avond beschikbaar waren en de deelnemers niet altijd geholpen werden met hun vraag of probleem. Deze nadelen zijn mogelijk ook aanwezig bij telefonisch contact, omdat beide opties door mensen worden beheerd gedurende specifieke tijden.

De opkomst van kunstmatige intelligentie (Artificial Intelligence: AI) biedt mogelijk een manier om de nadelen te voorkomen en de voordelen te behouden. AI wordt ten eerste gedefinieerd als een geprogrammeerde software die bedoeld is om de menselijke taal en het menselijk gedrag te simuleren, tijdens de omgang met mensen (Lei et al., 2021). In de klantenservice wordt AI voornamelijk ingezet om hulp te bieden aan klanten die contact opnemen met vragen, klachten of problemen. De ontwikkelingen in de technologie bieden bedrijven de mogelijkheid om AI-software in te zetten, om een deel van de klantenservice over te nemen. Zo wordt AI steeds vaker ingezet in de vorm van *chatbots* om menselijke chat agents te vervangen tijdens live chats (Morgan, 2023). Voor veel bedrijven is dit een economische maatregel, omdat AI op grote schaal goedkoper is dan menselijke medewerkers (Morgan, 2023).

Het inzetten van een chatbot in de klantenservice kan voordelen met zich meebrengen, zoals aangetoond in het onderzoek van Ho (2021). In tegenstelling tot een menselijke chat agent, wordt de responstijd van de live chat vrijwel per direct, kan de chatbot

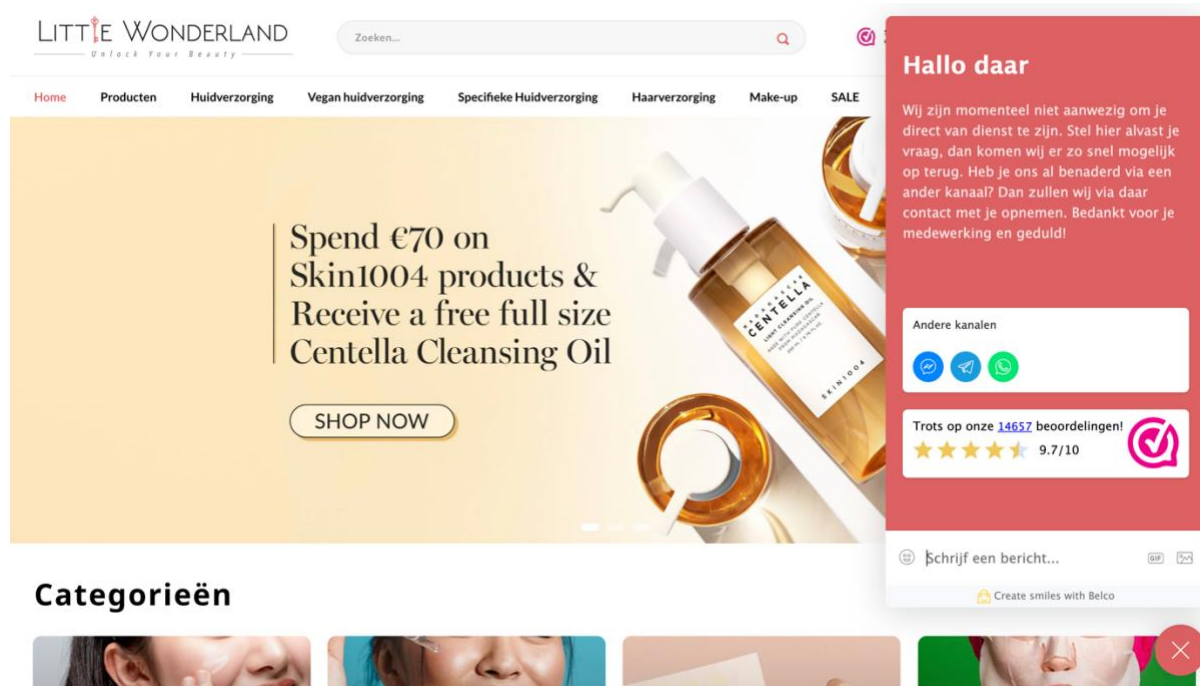
meerdere klanten tegelijkertijd helpen en kan de chatbot ook 24/7 actief zijn. Er zijn echter wel enkele nadelen aanwezig, zoals beschreven in het onderzoek van Dias en Correia (2021). Chatbots kunnen bijvoorbeeld met de huidige technologie nog niet compleet autonoom hun werk doen. Daarom moeten menselijke medewerkers de chat altijd in de gaten houden en moeten zij ook altijd op standby staan, in het geval dat de chatbot het gesprek niet goed kan beheersen en er een conflict ontstaat. Het onderzoek van Yang et al. (2020) heeft aangetoond dat ondanks de hoge competentie van chatbots, AI-chatbots wel fouten blijven maken tijdens de interactie met een mens. Fouten zoals het niet herkennen van de vragen of antwoorden, of meer technische problemen zoals het niet kunnen formuleren van antwoorden.

Onderzoek heeft zich tot nu toe vooral gericht op menselijke chat agents of chatbots, onafhankelijk van elkaar. Studies die beide onderwerpen met elkaar vergelijken zijn schaars. Vanwege het gebrek aan onderzoek is het moeilijk om duidelijke conclusies te kunnen trekken over de vergelijking tussen beide onderwerpen. Dit onderzoek heeft als doel om de ervaringen van mensen te vergelijken tussen de omgang met een chatbot en een menselijke chat agent. Door beide onderwerpen met elkaar te vergelijken, kan een duidelijker antwoord gegeven worden over de voorkeuren van mensen. De resultaten kunnen daarmee bijdragen aan toekomstige vervolgonderzoeken, die ook chatbots met menselijke medewerkers vergelijken. Dit onderzoeksgebied is namelijk nog in ontwikkeling, en dus niet uitgebreid onderzocht. Bovendien kunnen de bevindingen bedrijven in de toekomst ondersteunen bij het aanpassen van hun beleid omtrent chatbots. De onderzoeksvraag luidt: “Hoe ervaren personen de omgang met een AI-chatbot vergeleken met een menselijke chat agent in de klantenservice?”

2. Theoretisch kader

2.1 Live chats en de behoeftes van mensen

In het verleden werd klantenservice op afstand altijd voornamelijk gekenmerkt door een telefonisch gesprek of contact via e-mail. Awasthi (2020) laat zien dat dit de laatste jaren aan het afnemen is. De nieuw opkomende manier van klantenservice op afstand betreft de *live chats*. In 2022 koos bijvoorbeeld meer dan 85% van de bedrijven wereldwijd de optie om een live chat op hun website te implementeren (Awasthi, 2020). Live chats worden gekenmerkt door een geprogrammeerd chatvenster dat verschijnt op de website van het desbetreffende bedrijf. De live chat kan spontaan verschijnen, maar kan ook zichtbaar worden nadat de bezoeker op de ‘live chat’-optie heeft geklikt. In de onderstaande afbeelding is te zien hoe een typisch ‘live chat’-venster eruitziet. Een typische live chat wordt bovendien gekenmerkt door het aantal reeds vooraf geprogrammeerde opties, die de bezoeker kan kiezen om sneller de situatie uit te leggen en geholpen te worden. Opmerkelijk is dat live chats die alleen een optie bieden om met een menselijke medewerker te spreken, niet altijd beschikbaar zijn. Het is daarbij vaak wel mogelijk voor bezoekers om hun contactgegevens achter te laten wanneer zij hun vraag stellen, wat een gelijkenis heeft met het versturen van een onlinebericht. Dit is te zien in de onderstaande afbeelding.



Afbeelding 1. Een weergave van de website LittleWonderland.com met de ‘live chat’-optie met een menselijke chat agent.

Bezoekers kunnen in de live chat vervolgens vragen stellen of klachten uiten aan de medewerker die de chat beheert. Traditioneel gezien zijn dit menselijke *chat agents*, die in een bepaald tijdsvenster de klant of bezoeker kunnen helpen met het beantwoorden van een vraag of het oplossen van een probleem (Sanglé-Ferrière en Voyer, 2019). In het onderzoek van Helms en Mayo (2008) werd beschreven wat voor verwachtingen klanten hebben wanneer ze de klantenservice benaderen van een bepaald bedrijf. Enkele verwachtingen waren vriendelijkheid van de medewerker, een vlotte en foutloze service en het gevoel van gehoord worden. Bovendien liet het onderzoek van Kinna (2005) zien dat participanten behoefte hadden aan een medewerker die snel het probleem oploste, efficiënt te werk ging, altijd beschikbaar was en empathie toonde. Dit is in overeenkomst met het eerder besproken onderzoek van Helms & Mayo (2008). In beide onderzoeken werd hier echter niet altijd aan voldaan. Klanten hadden in het algemeen ook negatievere ervaringen, maar waren niet altijd open over wat zij hadden ervaren (Helms & Mayo, 2008).

Hoewel live chats een nieuw alternatief zijn voor de traditionele vorm van telefonisch contact, is het telefonisch contact niet geheel verdwenen uit de klantenservice. Live chats hebben echter wel duidelijke voordelen tegenover het telefonisch contact. Zoals beschreven in het onderzoek van Dias en Correia (2021) zijn klanten vaak ontevreden in het telefonisch contact, bijvoorbeeld omdat ze het bedrijf meerdere keren moeten contacteren voordat het probleem opgelost kan worden, of vanwege lange wachttijden, meerdere keren hetzelfde moeten herhalen, onvriendelijke medewerkers of medewerkers die het probleem uiteindelijk niet kunnen oplossen. Hier tegenover staan live chats, die bijvoorbeeld vrijwel geen wachttijden hebben, waarbij het gesprek lang door kan gaan en de medewerker alle berichten terug kan lezen, in het geval er iets mis is. Bovendien hebben live chats een bijzonder voordeel, namelijk dat gebruikers bijlagen kunnen versturen die de medewerker direct kan bekijken, om beter geholpen te worden. Concluderend hebben gebruikers van beide varianten dezelfde verwachtingen met betrekking tot de manier waarop zij willen dat het contact verloopt, maar toch zijn er in de praktijk wel verschillen in het proces van live chats tegenover telefonisch contact.

2.2 Kunstmatige Intelligentie als een klantenservice-chatbot

Een groot deel van de live chats wordt nu ook deels beheerd door een AI-chatbot (Morgan, 2023). Er zijn meerdere manieren waarop bedrijven hiermee omgaan. Een bedrijf kan bijvoorbeeld kiezen om eerst een chatbot in te zetten, de bezoeker zijn of haar vraag te

laten stellen en vervolgens een menselijke medewerker te laten komen. Het is ook mogelijk dat de bezoeker helemaal geen optie krijgt om met een menselijke medewerker te spreken, of alleen als de bezoeker hier meerdere keren nadrukkelijk om verzoekt. Bovendien is het op het eerste oog voor veel bezoekers niet duidelijk of de live chat op dat moment wordt beheerd door een menselijke chat agent of een chatbot. Live chats kunnen ook spontaan verschijnen en het begingesprek is in de meeste gevallen automatisch gegenereerd. In de onderstaande afbeelding is een typische live chat afgebeeld, die automatisch wordt aangeboden door een AI-chatbot. Mogelijk kan dit voor sommige bezoekers ook zorgen voor onzekerheid om het gesprek aan te gaan (Sanglé-Ferrière & Voyer, 2019).

AI-chatbots worden gekenmerkt door enkele belangrijke eigenschappen. Zoals beschreven in het onderzoek van Araujo (2018) speelt antropomorfisme een grote rol in hoe succesvol een AI-chatbot kan zijn. Antropomorfisme is een term die gebruikt wordt om de mate van menselijkheid van computers aan te geven. Computers krijgen eigenschappen toegewezen, die oorspronkelijk worden geassocieerd met mensen. Dit hangt nauw samen met het ‘Computers are Social Actors’-principe (CASA). Dit principe houdt in dat gebruikers in de omgang met computersystemen, dezelfde sociale gedragsregels eisen van deze systemen als tijdens de omgang met een medemens (Nass en Moon, 2000). Gebruikers gaan bijvoorbeeld praten tegen de computer, de computer begroeten of zelfs de computer uitschelden wanneer een probleem niet wordt opgelost. Hoewel het onderzoek van Nass en Moon (2000) niet onderzocht wat voor rol antropomorfisme speelde bij dit fenomeen, deed het onderzoek van Gambino et al. (2020) dat wel. Uit dat onderzoek bleek dat het CASA-principe het meest werd geobserveerd in de gevallen waarbij het computersysteem antropomorfische eigenschappen had.

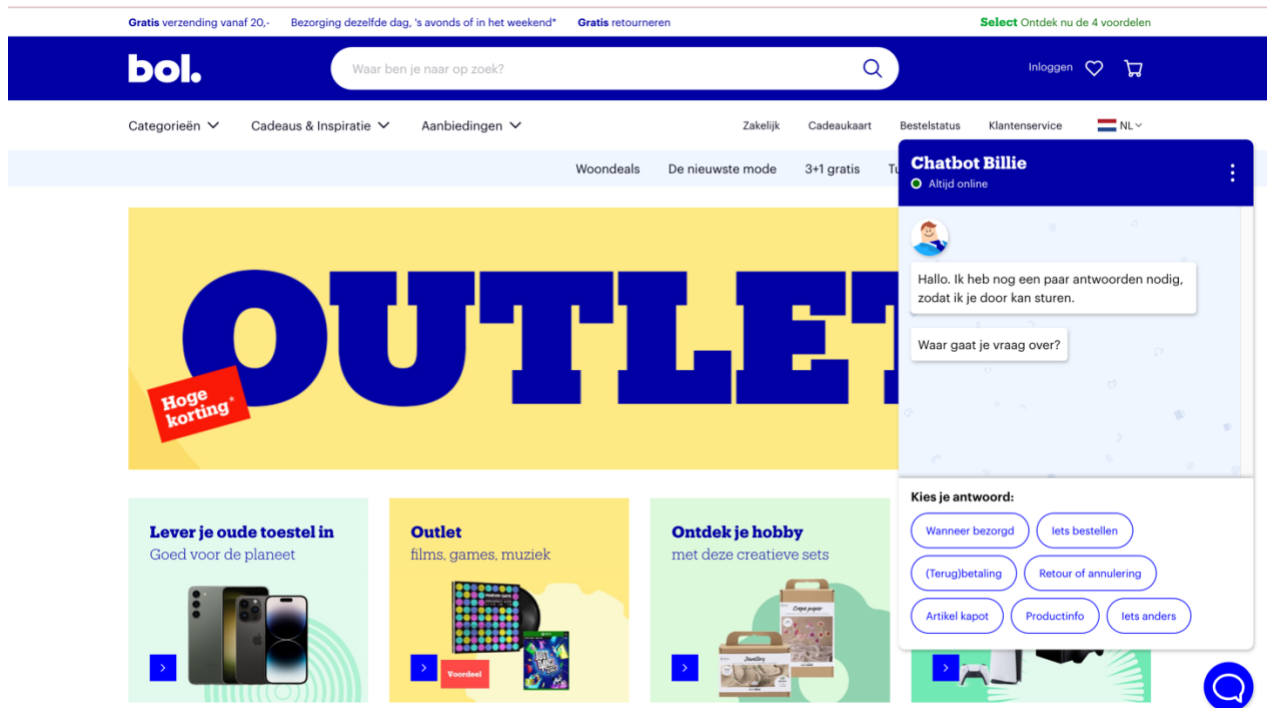
Er bestaan verschillende soorten antropomorfische chatbots. Araujo (2018) onderscheidt hierbij twee soorten: *disembodied* en *embodied*. *Disembodied* verwijst naar chatbots die geen virtuele avatar of lichaam hebben, maar slechts een chatprogramma zijn. *Embodied* verwijst naar chatbots die wel een avatar toegewezen hebben gekregen, zoals chatbot Billie van Bol.com (Afbeelding 2). De mate van antropomorfisme verschilt tussen de twee types. Bij het disembodied-type kan de mate van menselijkheid alleen geobserveerd worden in hoe de chatbot spreekt, klinkt of met de menselijke klant omgaat. Dit kan ook geobserveerd worden door middel van tekst, en niet alleen spraak. Bij het embodied-type daarentegen kan de mate van menselijkheid ook geobserveerd worden aan hoe menselijk de chatbot eruitziet. Daarnaast kan antropomorfisme een grote rol spelen in hoe tevreden klanten zijn met de service van een chatbot, zoals is aangetoond in het onderzoek van Wessel en

Benlian (2020). Zij benoemen dat antropomorfisme in een chatbot ervoor kan zorgen dat klanten vaker tevreden zijn en ook vaker hun feedback willen geven op de service. Dit geldt wel enkel voor de gevallen waarin de interactie succesvol verloopt en de chatbot voldoet aan de verwachtingen van de gebruiker. Antropomorfische kenmerken zorgen namelijk voor hogere verwachtingen bij de gebruiker, omdat bij de gebruiker het beeld ontstaat dat de chatbot een service kan bieden die vergelijkbaar is met de service die een menselijke medewerker kan verlenen (Wessel en Benlian, 2020).

Taal speelt een belangrijke rol in de mate van antropomorfisme van een chatbot. Gebruikers beoordelen over het algemeen dat de interactie met een chatbot vaak onnatuurlijk en onpersoonlijk oogt, door het taalgebruik van de chatbot (Liebrecht en Van Hooijdonk, 2020). Een menselijke manier van gespreksvoering tussen een chatbot en een gebruiker kan ervoor zorgen dat de gebruiker de chatbot menselijker beoordeelt en de interactie positief ervaart. Volgens het onderzoek van Liebrecht en Van Hooijdonk (2020) zijn er drie belangrijke maatregelen om ervoor te zorgen dat de chatbot op een menselijke en vloeiende manier een gesprek kan voeren met een gebruiker. Ten eerste bevelen ze aan dat de chatbot gebruikmaakt van gepersonaliseerd taalgebruik. Ten tweede bevelen ze aan dat de chatbot gebruikmaakt van begroetingen. Tot slot bevelen ze ook aan dat de chatbot zich aanpast aan het taalgebruik van de gebruiker. Door deze maatregelen toe te passen, krijgt de gebruiker meer zekerheid in de hulpverlening van de chatbot en voldoet de chatbot meer aan de verwachtingen die de gebruiker heeft (Liebrecht en van Hooijdonk, 2020).

Araujo (2018) noemt ook een tweede belangrijke component: *Social Presence*, oftewel sociale aanwezigheid. Dit is het gevoel dat er een ander levend of virtueel wezen samen met jou bestaat en op jou kan reageren. Een AI die zorgt voor een hoge mate van social presence, kan mogelijk aangenamer aanvoelen en de klant het gevoel geven dat de chatbot echt luistert en antwoord geeft. Aurajo (2018) benoemt dat deze twee factoren belangrijk zijn in de omgang tussen een mens en een AI-chatbot. Net als Aurajo (2018) benoemen Wessel en Benlian (2020) dat antropomorfisme kan zorgen voor een groter gevoel van sociale aanwezigheid, en daardoor ook voor meer tevredenheid bij klanten. Hieruit volgt de eerste hypothese: *“Mensen zullen een voorkeur hebben voor embodied chatbots met een hoge mate van antropomorfisme en social presence, vergeleken met disembodied chatbots.”* De categorie chatbots bestaat uit verschillende soorten, met verschillende gradaties van menselijkheid. Hoewel deze hypothese enkel twee verschillende chatbots vergelijkt en geen rekening houdt met menselijke medewerkers, kan het beantwoorden van deze hypothese een beter begrip bieden van de kenmerken die mensen belangrijk vinden in AI-chatbots en van de

wijze waarop deze kenmerken zich verhouden tot de eigenschappen van menselijke medewerkers. Dit kan hierdoor ook helpen om een genuanceerder antwoord te geven op de centrale onderzoeksvraag.



Afbeelding 2. Een weergave van de website Bol.com met de optie voor een live chat met een chatbot.

2.3 De verschillen tussen AI en een menselijke chat agent

Naast overeenkomsten, toont de literatuur aan dat er niet alleen belangrijke verschillen aanwezig zijn tussen een menselijke agent en een chatbot als fenomeen op zich, maar ook verschillen in hoe mensen de twee ervaren. Het onderzoek van Lei (2021) toonde bijvoorbeeld aan dat bij participanten vertrouwen een grote rol speelde in het kiezen voor een menselijke chat agent of een chatbot. Participanten hadden over het algemeen een hoger vertrouwen in menselijke medewerkers dan in chatbots. Daarnaast bleek uit de resultaten dat participanten een hogere score zouden geven aan hun algehele ervaring met een menselijke chat agent, dan aan de ervaring met een chatbot. Het onderzoek van Li en Zhang (2023) paste het *Push-Pull and Mooring Framework* (PPM) toe in de context van AI-chatbots in de klantenservice, vergeleken met een menselijke chat agent. Het PPM Framework houdt

rekening met push-, pull- en mooring-factoren die een rol spelen in de intentie van gebruikers om een keuze te maken tussen twee opties. Het onderzoek noemde push-effecten als factoren die gebruikers ontmoedigen om te kiezen voor een menselijke chat agent. Dit zijn factoren zoals een beperkt empathisch vermogen van de menselijke medewerker of een lage motivatie om zich aan te passen. Daarentegen verwijzen pull-effecten naar factoren die gebruikers juist aanmoedigen om te kiezen voor een AI-chatbot. Hierbij gaat het om factoren zoals rond-de-klok-beschikbaarheid en snelheid. Bovendien zijn mooring-effecten de factoren die een rol spelen bij de keuze die gebruikers maken. Dit betreft factoren zoals de regelmaat waarmee de gebruiker gebruikmaakt van de klantenservice en hoe belangrijk de gebruiker de klantenservice vindt. Hieruit volgt hypothese 2: *Mensen zullen een voorkeur hebben voor AI-chatbots vergeleken met menselijke chat agents, als gevolg van de push- en pull-factoren die aanwezig zijn.*

3. Methode

3.1 Databases en selectiecriteria

Om voldoende literatuur te vinden voor het beantwoorden van de onderzoeksvraag is gebruikgemaakt van de databases Google Scholar, Consensus, Web of Science en ACM. In alle vier databases zijn dezelfde relevante zoektermen gebruikt om geschikte literatuurartikelen te vinden. In de eerste instantie werd gezocht naar artikelen die een AI-chatbot vergeleken met een menselijke chat agent. Door een schaarste aan deze onderzoeken, werden ook artikelen geselecteerd die een AI-chatbot onderzochten of een menselijke chat agent, onafhankelijk van elkaar. Als artikelen andere vakgebieden behandelden, maar wel relevant waren voor het onderzoek, werden ze ook geselecteerd. Deze onderzoeken moesten daarbij wel chatbots met mensen vergelijken, of chatbots onafhankelijk van mensen, om ze te categoriseren als relevant voor het huidige onderzoek. Bovendien werd ernaar gestreefd om zoveel mogelijk artikelen te selecteren waarin gebruik werd gemaakt van experimentele onderzoeken.

Voor het selecteren van relevante artikelen, werd gekeken naar het gehele onderzoek, en niet slechts naar de samenvattingen. De gevonden artikelen zijn handmatig gelezen en gescand op hun relevantie. Dit werd gedaan om enkel de meest relevante artikelen te selecteren, die elementen bevatten die aansluiten op dit onderzoek. Ten eerste zijn artikelen geselecteerd met een experimenteel onderzoek, waarbij duidelijke resultaten aanwezig waren die meegenomen konden worden in dit huidige onderzoek. Ter aanvulling werden ook enkele kwalitatieve onderzoeken geselecteerd, die meer observerend te werk gingen. Bovendien waren er geen zeer specifieke selectiecriteria, zoals een specifiek land of jaartal, omdat dit niet relevant werd geacht. Tot slot is gekozen om uitsluitend Engelstalige artikelen te selecteren, om de gebruikte termen en uitleg zo consistent mogelijk te houden. Daarnaast zorgt dit voor een betere repliceerbaarheid van dit onderzoek, voor vervolgonderzoeken in de toekomst.

3.2 Zoektermen en overige zoekstrategieën

Een gecombineerde zoekterm werd gebruikt om een zo breed mogelijke resultatensectie te verkrijgen. De eerste zoekterm luidde: (*chatbot* OR "*AI*" OR "*artificial intelligence*") AND ("*customer service*" OR *webcare*) AND (*human* OR "*chat operator*" OR "*chat agent*"). Hiermee werden 1475 resultaten verkregen op de database ACM, 850 via

Consensus en 20.400 resultaten via Google Scholar. De zoekopdrachten werden op verschillende momenten bij alle vier de databases uitgevoerd, in de periode van 10 tot 22 maart 2024. Tot slot werd een techniek genaamd *snowballing* toegepast, waarbij relevante artikelen werden geselecteerd uit de bronverwijzingen van de gevonden artikelen.

3.3 Data-analyse

Uit 110 gescande artikelen, zijn er uiteindelijk vijftien geselecteerd voor dit onderzoek. De vijftien geselecteerde artikelen zijn overzichtelijk weergegeven in een tabel (zie tabel 1). Uit elk artikel zijn vervolgens zeven eigenschappen verzameld, die zijn gecategoriseerd onder de zeven kolommen in de tabel. De categorieën dienden daarnaast als hulpmiddel om ervoor te zorgen dat de artikelen kwalitatief goed waren en op elkaar waren afgestemd. Ten eerste is per artikel gekeken naar het publicatiejaar en het land van herkomst. Het was belangrijk dat alle geselecteerde artikelen actueel waren wat betreft de recente ontwikkelingen op AI-gebied. Er werd nagegaan of alle artikelen zo recent mogelijk waren gepubliceerd, of ten minste niet langer dan acht jaar geleden. Hier werd voor gekozen om enkel actuele informatie over te houden, omdat het vakgebied van kunstmatige intelligentie in de klantenservice in een snel tempo ontwikkelt. Het land van herkomst zorgde voor een overzicht van de verschillende landen waarin de onderzoeken hadden plaatsgevonden. Er zijn niet enkel onderzoeken uit Engelssprekende landen geselecteerd, waardoor de bevindingen gegeneraliseerd kunnen worden voor een internationale bevolking. Ten tweede werd gekeken naar het aantal participanten en het onderzoeksontwerp. Het was erg belangrijk om artikelen te selecteren die een groot aantal participanten hadden, waarbij de effecten ook makkelijker gegeneraliseerd konden worden. Bovendien werd ook gecategoriseerd wat de onderzochte variabelen waren, en daarbij of enkel de chatbots onderzocht werden of dat chatbots vergeleken werden met menselijke medewerkers. Daarnaast werd ook gecategoriseerd aan welke van de twee hypotheses het artikel zou bijdragen. Tot slot werd per artikel een korte samenvatting gegeven van de resultaten.

4. Resultaten

Om te onderzoeken hoe mensen de omgang met een chatbot ervaren in de klantenservice, vergeleken met een menselijke chat agent, zijn vijftien artikelen geselecteerd en onderzocht (zie tabel 1). Per artikel is ten eerste gekeken of er een voorkeur was voor chatbots of voor een menselijke chat agent. Ten tweede is gekeken welke elementen een rol speelden bij de gebruikers tijdens het bepalen van hun voorkeur. Bij de artikelen die enkel chatbots onderzochten, is gekeken naar het onderscheid tussen verschillende soorten chatbots, het type chatbot waarvoor de gebruikers de grootste voorkeur hadden, en hoe dit type zich verhoudt tot een menselijke medewerker.

Van de vijftien geselecteerde artikelen waren er acht die meerdere condities bevatten en zeven die slechts één conditie hadden. Daarnaast waren er negen artikelen die alleen chatbots onderzochten, en zes artikelen die chatbots met menselijke chat agents vergeleken. Bovendien kwamen vijf onderzoeken uit Noord-Amerika, zeven uit Azië, twee uit Europa en één uit Oceanië. Er waren zeven onderzoeken die meer dan vierhonderd participanten onderzochten, zeven onderzoeken die meer dan honderd participanten hadden en één onderzoek dat minder dan honderd participanten had. Daarenboven kregen participanten uit vier onderzoeken bestaande chatgesprekken toegewezen, waarbij ze zich moesten inbeelden dat zij het gesprek voerden. Participanten uit vier andere onderzoeken kregen de kans om zelf gesprekken te voeren en interactief bezig te zijn. Dit leverde in totaal acht artikelen op waarin participanten chatgesprekken moesten beoordelen. De overige zeven artikelen bevatten enkel enquêtes, waarvan twee artikelen een aanvullend onderzoek uitvoerden. Yen en Chang (2020) onderzochten daarbij ook de hersenactiviteiten van participanten, naast een enquête. Bovendien bevatte het onderzoek van Li en Zhang (2023) ook afgenomen interviews met enkele participanten, naast een enquête.

4.1 Embodied en disembodied chatbots

De eerste hypothese luidde: “Mensen zullen een voorkeur hebben voor embodied chatbots met een hoge mate van antropomorfisme en social presence, vergeleken met disembodied chatbots.” Tien van de vijftien behandelde artikelen hadden relevante bevindingen om een antwoord te geven op deze hypothese. Zes artikelen van de negen hadden positieve resultaten, die aansluiten op de eerste hypothese. Twee artikelen hadden deels ondersteunende resultaten, en twee artikelen vonden geen effect van een hoge mate van antropomorfisme en social presence bij chatbots op de voorkeur van mensen (De Cicco et al.,

2020). Tot slot hadden elf artikelen wel betrekking op de klantenservice en vier artikelen niet. In het onderzoek van Nguyen et al. (2023) moesten participanten middels een enquête stellingen beantwoorden over het empathisch vermogen van een chatbot, de anonimiteit en personalisatie die chatbots bieden, en de mate waarin er vertrouwen is in chatbots. De bevindingen toonden aan dat gebruikers een sterke voorkeur hadden voor een chatbot die menselijke eigenschappen vertoonde, zoals een empathisch vermogen, en luisterde naar de wensen van de gebruikers en daarop de service aanpaste.

Deze bevindingen waren in overeenkomst met de resultaten van het onderzoek van Roy en Naidoo (2020). In dit onderzoek werd onderscheid gemaakt tussen twee typen chatbots: chatbots met een ‘warme’ persoonlijkheid en ‘vaardige’ chatbots, die competent waren in hun werkzaamheden, maar geen duidelijke persoonlijkheid hadden. Een warme persoonlijkheid werd bijvoorbeeld gedefinieerd als het hebben van een empathisch vermogen, een vriendelijke manier van praten en gewilligheid om problemen succesvol op te lossen. Participanten kregen bestaande chatgesprekken te zien tussen een klant en een chatbot, waarbij ze zich moesten inbeelden dat zij de klant waren. De participanten werden geplaatst in twee condities, waarbij ze chatgesprekken kregen van een chatbot met een ‘warme’ persoonlijkheid of gesprekken met een ‘vaardige’ chatbot. Participanten moesten vervolgens stellingen beantwoorden over hun mening met betrekking tot het type chatbot en hun gewilligheid om middels de chatbot een aankoop te doen. De bevindingen toonden aan dat gebruikers een sterke voorkeur hadden voor chatbots met menselijke, ‘antropomorfische’ kenmerken, in combinatie met een ‘warme’ persoonlijkheid.

De resultaten in het onderzoek van Yen en Chiang (2020) toonden iets vergelijkbaars. De participanten werden geplaatst in twee condities, waarbij ze een online-winkelpagina met of zonder een chatbotfunctie moesten beoordelen. Participanten moesten vervolgens stellingen over chatbots beantwoorden, gebaseerd op verschillende constructen zoals vertrouwen, antropomorfisme, vaardigheden en sociale aanwezigheid. Daarbij werd hun hersenactiviteit ook onderzocht door middel van een hersenfilmpje. Participanten kregen meer vertrouwen in chatbots als zij antropomorfische kenmerken vertoonden, vaardig waren in hun werkzaamheden en getuigden van sociale aanwezigheid. De participanten hadden daarbij een sterke voorkeur voor deze soorten chatbots, tegenover chatbots die deze eigenschappen niet bezaten.

De resultaten van het onderzoek van Schuetzler et al. (2020) waren in overeenkomst met de eerder genoemde onderzoeken. In dit onderzoek werden participanten onderverdeeld in vier condities, met vier verschillende typen chatbots. De chatbots varieerden wat betreft de

mate waarin zij het gesprek personaliseerden naar de voorkeuren van de gebruiker, en hoe gevarieerd hun spraak was. De participanten moesten eerst interactieve gesprekken voeren met de chatbot. Vervolgens moesten zij stellingen beantwoorden gebaseerd op hoe menselijk de chatbot overkwam, hoe interactief de chatbot was en in hoeverre de chatbot sociaal aanwezig was. De resultaten lieten zien dat de participanten een voorkeur hadden voor chatbots die meer antropomorfische eigenschappen vertoonden en een grotere sociale aanwezigheid hadden. Dit soort chatbots kwam vervolgens vaardiger en kompetenter over, vergeleken met chatbots die minder menselijke kenmerken vertoonden.

Het onderzoek van Ashfaq et al. (2020) ondersteunde de resultaten van de eerder behandelde onderzoeken. Participanten moesten in het onderzoek stellingen beantwoorden gebaseerd op hun ervaringen met chatbots in de klantenservice, over meerdere verschillende constructen, zoals hoe informatief zij een chatbot vonden, hoe kwalitatief zij de service vonden, hoe aangenaam hun ervaringen waren, hoe nuttig zij de omgang met een chatbot vonden en hoe makkelijk zij de omgang met een chatbot vonden, en met name hoe menselijk en natuurlijk de gespreksvoering verliep. Daarnaast werd ook onderzoek gedaan naar tevredenheid, te weten in hoeverre zij verder zouden willen gaan met het gebruik van een chatbot en in hoeverre zij de behoefte ondervonden om met een menselijke medewerker te spreken. De bevindingen toonden aan dat gebruikers die een positieve ervaring hadden met een chatbot en daarbij hoog scoorden op de constructen, geen significante behoefte hadden om tijdens de omgang met een chatbot een menselijke medewerker erbij te willen halen.

Het onderzoek van Svenningsson en Faraon (2019) resulteerde in enkele opmerkelijke bevindingen. Participanten moesten stellingen beantwoorden gebaseerd op hun mening over verschillende concepten, zoals in hoeverre de chatbot naast werkzaamheden ook informele gesprekken kon voeren, de manier waarop de chatbot de werkzaamheden zou moeten uitvoeren, wat voor taalgebruik (informeel of formeel) de chatbot zou moeten gebruiken, hoe de chatbot de gebruiker zou moeten aanspreken en in hoeverre de chatbot emotie zou moeten vertonen. Hierbij beeldden de participanten zich een hypothetisch gesprek in en hoe zij verwachtten dat een chatbot zich zou gedragen. De bevindingen toonden daarbij aan dat participanten een sterke voorkeur hadden voor chatbots die nadrukkelijk aangaven dat zij chatbots waren. Bovendien preferden participanten dat de chatbot minimale humor vertoonde en duidelijk was in het aanspreken. Er was echter wel een voorkeur voor chatbots die menselijke eigenschappen konden vertonen, zoals emoties. Tot slot kwam naar voren dat er sterke verschillen waren in hoe participanten gesprekken wilden voeren met chatbots, en dat daarbij personalisatie in het gesprek tussen mens en chatbot erg belangrijk zou zijn.

De bevindingen van het onderzoek van Brandtzæg en Følstad (2017) toonde ook opmerkelijke resultaten. In dit onderzoek moesten participanten hun meningen geven over het gebruik van chatbots, en aangeven voor welke doeleinden zij chatbots wilden gebruiken in plaats van menselijke chat agents. De participanten waren hierbij vrij om zelf hun antwoorden in te voeren, en werden niet getoetst op bepaalde vooraf bedachte constructen. De resultaten lieten zien dat in de meeste gevallen participanten een voorkeur hadden voor chatbots die sociale vaardigheden hadden en menselijk taalgebruik. Daarnaast gaven ze aan dat een chatbot met menselijke kenmerken meer hun behoeftes zou kunnen vervullen, dan een chatbot zonder deze kenmerken. Bovendien waren er verschillende redenen waarom gebruikers ervoor kozen om een chatbot te gebruiken. Een van de meest voorkomende redenen om een chatbot te gebruiken was om productiever te zijn en sneller taken af te krijgen, en om een sociale behoefte te vervullen. Een sociale behoefte werd beschreven als een behoefte om de aanwezigheid van een ander te voelen en samen met diegene in gesprek te gaan.

Het onderzoek van Gümüş en Çark (2021) bood ondersteunde resultaten aan alle voorafgaande onderzoeken. In dat onderzoek moesten participanten stellingen beantwoorden over het gebruik van chatbots. Er waren verschillende constructen, zoals hoe nuttig zij chatbots vonden, hoe makkelijk in gebruik chatbots waren, hun mate van tevredenheid over chatbots, de mate van gevaar of risico die chatbots meebrachten, hun algehele klantenervaring en hoe graag zij gebruik wilden maken van een chatbot. De bevindingen toonden aan dat hoe makkelijker in gebruik participanten een chatbot vonden, hoe hoger chatbots ook scoorden op de andere onderdelen. Bovendien speelde de mate van gevaar of risico bij chatbots geen rol in hoe laag of hoog de chatbots scoorden op andere onderdelen. Participanten vonden daarbij het risico van chatbots niet belangrijk genoeg om ze volledig af te schrijven.

Het onderzoek van De Cicco et al. (2020) had tegenstrijdige resultaten met alle voorafgaande onderzoeken. In het onderzoek werden participanten onderverdeeld in vier condities, met elk een chatbot. De chatbots waren ofwel gericht op sociale interactie, met of zonder een avatar, of ze waren gericht op het vervullen van taken, met of zonder een avatar. De participanten moesten in alle vier de condities een interactief gesprek voeren met de desbetreffende chatbot. Vervolgens moesten de participanten stellingen beantwoorden over hun ervaringen met de chatbot, met constructen zoals de mate van sociale aanwezigheid, tevredenheid, vertrouwen en hun algemene mening over de chatbot. De resultaten lieten zien dat participanten niet een duidelijke voorkeur lieten zien voor embodied chatbots, die sociaal

interactief waren, tegenover disembodied chatbots die slechts taken konden uitvoeren. Het bleek in het onderzoek dat antropomorfische kenmerken niet genoeg waren om een duidelijke voorkeur te vormen.

Tot slot had het onderzoek van Song et al. (2022) ook een resultaat dat tegenstrijdig is met de hypothese. Hoewel het onderzoek keek naar de voorkeuren van mensen met betrekking tot chatbots en menselijke chat agents, had het een opmerkelijke bevinding. Uit de resultaten bleek dat de menselijke eigenschappen van een chatbot, oftewel het antropomorfisme, ook een rol speelden in hun voorkeur. In sommige gevallen werden participanten juist ongemakkelijker door de menselijke kenmerken van een chatbot.

4.2 AI chatbots en menselijke chat agents

De tweede hypothese luidde: “Mensen zullen een voorkeur hebben voor AI-chatbots vergeleken met menselijke chat agents, door de push- en pull-factoren die aanwezig zijn.” Van de vijftien behandelde artikelen hadden zes artikelen relevante bevindingen om een antwoord te geven op de tweede hypothese. Eén van de zes artikelen had volledig ondersteunende resultaten. Daarnaast waren twee artikelen deels ondersteunend en deels niet. Tot slot hadden drie artikelen resultaten die de hypothese tegenspreken: deze toonden juist aan dat mensen een voorkeur hadden voor menselijke chat agents en niet voor AI-chatbots.

Het onderzoek van Li en Zhang (2023) ondersteunde de hypothese. In het onderzoek werden participanten geïnterviewd over hun mening met betrekking tot chatbots en menselijke chat agents. Vervolgens deden ze mee aan een enquête. Zij werden getoetst op meerdere constructen, zoals hoe empatisch zij een chatbot of menselijke chat agent vonden, hoe verbonden zij zich met hen voelden, en in hoeverre de chatbot of menselijke chat agent zich goed konden aanpassen aan het gesprek en aan de voorkeuren van de gebruiker. De participanten werden geïnstrueerd om te antwoorden op basis van hun eerdere ervaringen. De resultaten toonden aan dat participanten chatbots een hogere score gaven, door de aanwezige push- en pull-effecten. Chatbots die bijvoorbeeld empathischer overkwamen en beter het gesprek konden personaliseren, creëerden ‘pull-effecten’ bij de gebruiker om een voorkeur te vormen voor chatbots tegenover menselijke chat agents. Bovendien zorgden menselijke chat agents die bijvoorbeeld minder behulpzaam en vriendelijk overkwamen voor een ‘push-effect’ om de gebruiker juist over te laten gaan op chatbots. Daarnaast werd ook aangetoond dat de gebruikers die aangaven ervaren te zijn in de klantenservice, vaker een sterke voorkeur hadden voor chatbots dan menselijke chat agents.

De bevindingen van het onderzoek van Chen et al. (2023) waren deels ondersteunend. Ten eerste werden de participanten onderverdeeld in vier condities. Per conditie moesten participanten een chatbot of een menselijke chat agent beoordelen, in een situatie met serviceproducten (*experience products*) of tastbare producten (*search products*). Bovendien moesten participanten reeds bestaande chats beoordelen, in alle vier de condities. Bij de beoordeling moesten de participanten stellingen beantwoorden over hun aankoopintentie, hoe vlot zij de interactie vonden verlopen en hoe kwalitatief zij de klantervaring vonden. De resultaten toonden aan dat participanten chatbots een hogere score gaven in de situatie waarin zij zochten naar tastbare producten. In de situatie met serviceproducten kregen menselijke chat agents juist een hogere score, vergeleken met chatbots. Het blijkt uit het onderzoek dat er een verschil is of er voor een chatbot of een menselijke chat agent wordt gekozen, afhankelijk van het type product. Dit ondersteunt de hypothese deels.

Het onderzoek van Folk et al. (2022) toonde noemenswaardige bevindingen. In dit onderzoek werden participanten onderverdeeld in vier condities. Participanten kregen te maken met een chatbot of een menselijke medewerker, in een sociale setting of een minder sociale setting. De participanten moesten eerst een gesprek voeren met de chatbot of menselijke chat agent, en vervolgens beoordelen hoe zij vonden dat het gesprek ging. De participanten kregen te horen dat zij een gesprek gingen voeren met een medeparticipant. Zij wisten niet dat zij een gesprek zouden voeren met een menselijke chat agent, of een chatbot. De beoordeling werd gegeven aan de hand van twee constructen: de mate van sociale connectie en hun algemene tevredenheid. Uit de resultaten bleek dat participanten een sterkere voorkeur hadden voor sociale chatbots, vergeleken met minder sociale menselijke chat agents. Wanneer de menselijke chat agents echter sociaal waren, werd voor hen gekozen. De voorkeur van de participanten was daarmee afhankelijk van de situatie, wat de hypothese deels ondersteunt.

Het onderzoek van Folk et al. (2022) was in tegenstelling met de voorafgaande onderzoeken. Participanten werden nogmaals in vier condities geplaatst, waarbij ze bestaande gesprekken met een chatbot of een menselijke chat agent moesten lezen en beoordelen. Daarnaast waren er twee typen gesprekssituaties, een informatief *education script* of een amuserend *entertainment script*. Beide gesprekssituaties gingen over een aankoop in een winkel. Vervolgens moesten de participanten stellingen beantwoorden over hun tevredenheid en hun aankoopintentie na het lezen van de gesprekken. Uit de resultaten bleek dat participanten een hogere score gaven aan menselijke chat agents in de informatieve

gesprekssituatie, vergeleken met chatbots. Bovendien was in alle gevallen de tevredenheid en aankoopintentie hoger bij menselijke chat agents, dan bij chatbots.

Het onderzoek van Lei et al. (2021) liet resultaten zien die de tweede hypothese tegenspreken. De participanten werden in twee condities geplaatst, waarbij ze stellingen moesten beantwoorden over hun ervaring met een chatbot, of hun ervaring met een menselijke chat agent. Daarbij werden meerdere constructen getoetst, zoals de mate van sociale aanwezigheid, hoe groot hun vertrouwen was, hoe fijn de interactie was op sociaal niveau en hoe geneigd ze zouden zijn om de gesprekspartner taken te laten uitvoeren. De resultaten lieten zien dat participanten een hogere score gaven aan menselijke chat agents op het niveau van vertrouwen en toegankelijkheid, vergeleken met chatbots.

Tot slot waren de bevindingen uit het onderzoek van Song et al. (2022) ook in tegenspraak met de hypothese. Participanten werden in twee condities geplaatst, waarbij zij bestaande gesprekken moesten beoordelen die gevoerd waren met een chatbot of met een menselijke chat agent. Vervolgens moesten zij stellingen beantwoorden over hun voorkeur, en of zij de intentie hadden om het gesprek zelf ook te gaan voeren met de chatbot of de menselijke chat agent. De bevindingen toonden dat participanten over het algemeen een sterkere voorkeur hadden voor menselijke chat agents, vergeleken met chatbots.

Tabel 1

Weergave van alle geselecteerde artikelen

Auteur	N	Public atie jaar	Land	Design	Variabelen (OV, AV)	Th em a	Gevonden resultaten
Lei et al.	400	2021	China	2 condities, tussenproef persoonont werp	OV: Chatbots VS menselijke medewerkers AV: “social	H2	Menselijke chat agents kregen een hogere score, voor

					presence”, “task attraction”, “trust” en “social attraction”		vertrouwen en toegankelijkheid, dan chatbots.
Li en Zhang	441	2023	China	1 conditie, binnenproef persoonontwerp	OV: Chatbots VS menselijke medewerkers AV: onder andere “empathy”, “connectivity”, “adaptability”, “personalization”	H2	Chatbots kregen een hogere score, door push & pull effecten. Ervaren gebruikers van klantenservice hadden een sterkere voorkeur voor chatbots dan menselijke agents.
De Cicco et al.	193	2020	Italië	2x2, tussenproef persoonontwerp	OV: Embodied VS non-embodied, social-oriented VS task-oriented	H1	Onderzoek alleen naar chatbots. Antropomorfe kenmerken waren niet voldoende om een

					AV: “social presence”, “enjoyment”, “trust” en “attitude”		voorkeur voor een bepaalde type chatbot te creëren, bijvoorbeeld chatbots zonder en met antropomor fische kenmerken.
Chen et al.	234	2023	China	2x2, tussenproef persoon twerp	OV: AI VS menselijke medewerker, search products VS experience products	H2	Chatbots kregen een hogere score wanneer gebruikers zochten naar gewone producten. Bij serviceprod ucten kregen menselijke agents een hogere score.
					AV: “purchase intention”, “processing fluency” en “perceived service quality”		

Sands et al.	404	2020	VS	2x2, tussenproef persoonontwerp	OV: chatbot VS menselijke medewerker, entertainment script VS education script AV: “satisfaction” en “purchase intention”	H2	Menselijke chat agents kregen een hogere score dan chatbots, wanneer zij een informatief gesprek voerden. Daarnaast was de aankoopintentie en tevredenheid hoger voor menselijke chat agents dan chatbots.
Nguyen et al.	413	2023	Vietnam	1 conditie, binnenproef persoonontwerp	OV: Chatbots AV: “empathy response”, “anonymity”, “customization”	H1	Onderzoek alleen naar chatbots. Personalisatie en een empathisch vermogen bij een chatbot, zorgde voor

					en “consumer trust”		een sterke voorkeur in vergelijking tot chatbots zonder deze eigenschappen.
Gümüş en Çark	211	2021	Turkije	1 conditie, binnenproef persoonontwerp	OV: Chatbots AV: “perceived usefulness”, “perceived ease of use”, “perceived enjoyment”, “perceived risk”, “customer experience” en “behavioral intention”	H1	Onderzoek alleen naar chatbots. Chatbots die gemakkelijk in gebruik waren, kregen een hoge score. Daarnaast speelde risico op een slechte ervaring, geen rol in de voorkeur naar chatbots.
Roy en Naidoo	446	2020	Australië	2 condities, tussenproef persoonontwerp	OV: “Warme” chatbots VS vaardige chatbots	H1	Onderzoek alleen naar chatbots. Gebruikers hadden een voorkeur

					AV: “attitude” en “purchase intention”		voor chatbots met antropomor fische kenmerken, gecombinee rd met een warme persoonlijk heid.
Folk et al.	401	2022	Canada	2x2, tussenproef persoonont werp	OV: Chatbots VS menselijke medewerkers, sociaal VS minder sociaal AV: “social connection” en “happiness”	H2	Onderzoek naar gespreksvo ering. Gebruikers hadden een sterke voorkeur voor sociale chatbots vergeleken met asociale menselijke agents.
Yen en Chiang	204	2020	China	1 conditie, binnenproef persoonont werp	OV: Chatbots AV:	H1	Onderzoek alleen naar chatbots. Gebruikers vonden dat

					“trust” en “purchase intention”		het vertrouwen in chatbots werd verbeterd door goede vaardighed en, antropomor fische kenmerken en een sociale aanwezighe id.
Brandtzæ g en Følstad	146	2017	VS	1 conditie, binnenproef persoonont werp	OV: Chatbots AV: Geen afhankelijke variabele, participanten waarin vrij om hun redenen zelf in te voeren.	H1	Onderzoek alleen naar chatbots. Gebruikers hadden een sterke voorkeur voor chatbots ten opzichte van mensen, om productieve r te zijn, maar deels ook voor sociale

							relaties.
Svenning sson en Faraon	75	2019	Zweden	1 conditie, binnenproef persoonont werp	OV: Chatbots AV: “small talk”, “way of conduct in tasks”, “sophisticated choices of words”, “dynamic approach” en “expressing feelings”	H1	Onderzoek alleen naar chatbots. Gebruikers hadden een voorkeur voor chatbots die zich identificeer den als chatbots, die een duidelijke spraak hadden met minimale humor. Daarnaast was er een voorkeur voor het vertonen van menselijke eigenschap pen zoals emoties. Bovendien was er een groot

							verschil in hoe gebruikers gesprekken voerden met chatbots.
Schuetzler et al.	450	2020	VS	4 condities, tussenproef persoonont werp	OV: Generic Varied VS Generic Non- Varied VS Tailored Varied VS Tailored Non- Varied AV: “humanness”, “engagement” en “social presence”	H1	Onderzoek alleen naar chatbots. Gebruikers hadden een voorkeur voor chatbots met goede vaardigheid en. Zulke chatbots vertoonden daardoor meer antropomor fische kenmerken en hadden een grotere sociale aanwezighe id.
Ashfaq et al.	370	2020	VS	1 conditie, binnenproef	OV: Chatbots	H1	Onderzoek alleen naar

				persoonont werp	AV: “information quality”, “service quality”, “perceived enjoyment”, “perceives usefulness”, “perceived ease of use”, “need for interaction with a service employee”, “satisfaction” en “continuance intention”		chatbots. Gebruikers die een positieve ervaring hadden met chatbots (en daarbij bijv. een hogere IQ, SQ of “perceived enjoyment” scoorden, hadden geen significante behoefte om met een menselijke medewerke r te spreken.
Song et al.	125	2022	China	2 condities, tussenproef persoonont werp	OV: Chatbots VS menselijke medewerkers AV: “Adoption intention”	H2 en H1	Gebruikers hadden over het algemeen een sterke voorkeur naar menselijke medewerke

rs, ten
opzichte
van
chatbots. In
sommige
gevallen
werkten
antropomor
fische
kenmerken
van een
chatbot
juist
averechts.
(Gebruikers
vonden het
ongemakke
lijk)

5. Discussie en conclusie

5.1 Discussie

Dit onderzoek had als centrale onderzoeksvraag: “Hoe ervaren personen de omgang met een AI-chatbot vergeleken met een menselijke chat agent in de klantenservice?” Om de onderzoeksvraag en de daaruit voortvloeiende hypothesen te beantwoorden, werd een literatuurstudie uitgevoerd waarbij vijftien relevante artikelen werden geselecteerd en bestudeerd.

De eerste hypothese in dit onderzoek ging over de voorkeur van mensen voor een *embodied chatbot* met een hoge mate van antropomorfisme en *social presence*, vergeleken met *disembodied chatbots*. Negen van de vijftien onderzochte artikelen keken naar dit aspect. In zes van de negen artikelen werd ondersteund dat antropomorfisme en sociale aanwezigheid een rol speelt bij mensen in het kiezen voor een chatbot. Twee artikelen ondersteunden dit deels en slechts een enkel artikel had geen bevindingen die dit ondersteunden. Aangezien slechts twee artikelen (De Cicco et al., 2020 en Song et al., 2022) de hypothese niet konden ondersteunen en de overige acht artikelen deze wel konden bevestigen, kan de hypothese aangenomen worden. Bovendien werd de hypothese bevestigd door eerdere onderzoeken, die hadden aangegeven dat antropomorfisme een belangrijke rol speelde in hoe positief gebruikers een chatbot ervoeren. Daarnaast gaven ze aan dat menselijke eigenschappen in chatbots ervoor zorgden dat gebruikers zich meer sociaal verbonden voelden met chatbots.

Het onderzoek van De Cicco et al. (2020) geeft mogelijke verklaringen voor het feit dat er geen duidelijke bevindingen waren dat antropomorfisme een rol speelde bij gebruikers om een voorkeur te vormen voor een bepaald type chatbot. In dit onderzoek mochten participanten zelf een gesprek voeren met een chatbot die gericht was op sociale interactie of met een chatbot die enkel gericht was op het uitvoeren van taken, in een conditie met een avatar of zonder avatar. Ten eerste is het mogelijk dat de mate van *embodiment*, wat een chatbot met en zonder avatar van elkaar onderscheidde, niet duidelijk genoeg overkwam op de participant, waardoor het effect ook niet naar voren zou kunnen komen. Daarnaast is het mogelijk dat doordat participanten zelf de mogelijkheid kregen om een gesprek te voeren met de chatbot, ze zelf ook een bepaalde mate van sturing gaven in hoe sociaal interactief het gesprek was, waardoor dat onderscheid ook niet duidelijk naar voren kwam. Bovendien gaven de auteurs van het onderzoek aan dat participanten in een fictieve situatie het gesprek moesten uitvoeren, wat mogelijk een rol zou kunnen spelen in de generaliseerbaarheid naar

werkelijke situaties. Tot slot is het mogelijk dat de tegenstrijdige bevindingen het gevolg zijn van de manier waarop antropomorfisme is geoperationaliseerd per onderzoek.

Antropomorfisme wordt bijvoorbeeld in enkele onderzoeken neergezet als iets visueels, met een avatar. In andere onderzoeken wordt antropomorfisme juist geassocieerd met het gedrag en taalgebruik van de chatbot. De verschillende operationalisaties kunnen ook een oorzaak vormen waarom antropomorfisme in sommige gevallen averechts blijkt te werken en mensen juist ongemakkelijk maakt. De verschillende definities kunnen ervoor gezorgd hebben dat er geen generaliseerbare conclusie getrokken kan worden.

De tweede hypothese in dit onderzoek ging over de voorkeur van mensen voor AI-chatbots vergeleken met menselijke chat agents, door de aanwezige push- en pull-factoren. Eerder onderzoek bevestigde het idee dat push-factoren zoals onvriendelijkheid van een menselijke medewerker en lange wachttijden, en daarnaast pull-factoren zoals de snelle service van een chatbot en personalisatie, een rol kunnen spelen in het vormen van een voorkeur voor chatbots vergeleken met menselijke chat agents. Van de vijftien artikelen waren zes artikelen relevant om de hypothese te kunnen toetsen. Drie artikelen ondersteunden de hypothese deels en twee artikelen hadden resultaten die volledig in tegenspraak waren (Lei et al., 2021 en Song et al., 2022). Slechts een enkel artikel had bevindingen die de hypothese konden ondersteunen (Li en Zhang, 2023). Er zijn mogelijk factoren die een rol spelen in het bepalen van de voorkeur van mensen voor een chatbot of een menselijke chat agent.

Voorbeelden hiervan zijn het type product of service dat wordt behandeld, of de mate van socialiteit van de chatbot of menselijke chat agent. Over het algemeen hadden participanten ook meer vertrouwen in menselijke medewerkers dan in chatbots. Dit kan ook komen doordat chatbots relatief nieuw zijn en gebruikers nog geen vertrouwen hebben opgebouwd in de hulpverlening van chatbots. Daarnaast zijn er ook verschillen in de behandelde artikelen. Ten eerste kregen participanten in het onderzoek van Lei et al. (2021) een enquête om in te vullen, waarin ze werden bevraagd over hun ervaringen met een chatbot vergeleken met een menselijke medewerker. De participanten kregen geen kans om interactieve gesprekken te voeren met chatbots of menselijke chat agents, waarbij resultaten met elkaar vergeleken konden worden, zoals wel gebeurde in de onderzoeken van Chen et al. (2022) en Folk et al. (2023). Het is mogelijk dat het gebruik van enkel een enquête op een andere manier de voorkeur van participanten ging meten. Daarbij is het mogelijk dat een andere onderzoeksmethode andere resultaten zou tonen. Ten tweede werden in het onderzoek van Song et al. (2022) bestaande geschreven gesprekken beoordeeld door de participant. Dit is vergelijkbaar met de methode van het onderzoek van Lei et al. (2021), omdat in feite enkel

naar de meningen werd gevraagd van de participanten. Het is mogelijk dat indien participanten zelf een gesprek zouden kunnen voeren, er andere bevindingen naar voren zouden komen. Vanwege de verschillende tegenstrijdige bevindingen, kan hypothese 2 niet aangenomen worden.

5.2 Theoretische en praktische implicaties

Hoewel er wel enkele studies blijken te zijn die duidelijke bevindingen hebben, zijn onderzoeken naar dit onderwerp schaars; met name de onderzoeken die chatbots met menselijke chat agents vergelijken met betrekking tot de klantenservice. Dit literatuuronderzoek is een van de weinige onderzoeken die een poging heeft gewaagd om beide met elkaar te vergelijken. Ten eerste kunnen de resultaten van deze studie meer inzicht bieden aan onderzoekers in toekomstige vervolgonderzoeken. Bovendien kunnen de bevindingen van dit onderzoek bijdragen aan het verkleinen van de kenniskloof met betrekking tot de voorkeuren van mensen voor een chatbot of een menselijke chat agent. Daarnaast kunnen de resultaten van dit onderzoek bedrijven en beleidsmakers helpen om beter te begrijpen onder welke omstandigheden chatbots of menselijke chat agents worden geprefereerd door gebruikers. Zij zullen daarbij beter geïnformeerd worden hoe zij de klantenservice kunnen indelen aan de hand van de behoeftes en verwachtingen van gebruikers. Tot slot kunnen deze onderzoeksresultaten andere vakgebieden meer inzicht bieden in de gevallen waarin een chatbot ingezet kan worden voor hulpverlening.

5.3 Limitaties en suggesties voor vervolgonderzoek

Er zijn enkele kanttekeningen die gemaakt kunnen worden bij de bevindingen van de behandelde artikelen, waar rekening mee gehouden kan worden voor toekomstige vervolgonderzoeken. Ten eerste werden slechts vijftien artikelen geselecteerd om de onderzoeksvraag te beantwoorden, hoewel het een vrij complexe onderzoeksvraag was. In vervolgonderzoeken kan gestreefd worden naar een grotere hoeveelheid artikelen, om een duidelijker beeld te krijgen van de bevindingen en de generaliseerbaarheid van de conclusie te vergroten. Daarbij werd in dit onderzoek een kleine hoeveelheid artikelen behandeld die chatbots vergeleek met menselijke chat agents in hetzelfde onderzoek. In het vervolg kan tijdens het selectieproces de nadruk gelegd worden op onderzoeken die beide vergelijken.

Voor vervolgonderzoeken zijn er enkele suggesties waar rekening mee gehouden kan worden met betrekking tot de operationalisatie. In het huidige onderzoek zijn er bijvoorbeeld

slechts enkele onderzoeken die participanten interactief een gesprek laten voeren met een chatbot of menselijke chat agent. De overgrote meerderheid van de behandelde artikelen bevatte alleen een enquête. Het voeren van een interactief gesprek in een experimentele conditie kan op een duidelijke manier meer inzicht geven in de directe voorkeuren van de gebruikers. Dit kan de generaliseerbaarheid en betrouwbaarheid van de conclusies ook vergroten. Tot slot werd participanten in sommige onderzoeken pas achteraf verteld dat zij in gesprek waren met een chatbot. In andere onderzoeken werden participanten echter vooraf verteld dat zij een gesprek zouden gaan voeren met een chatbot. Hierbij verschilt de kennis die participanten hebben in beide situaties, wat mogelijk gevolgen kan hebben voor de bevindingen van deze onderzoeken. In het vervolg kan rekening gehouden worden met de potentiële verschillen die beide onderzoeksmethodes kunnen veroorzaken in de resultaten.

5.4 Conclusie

In dit literatuuronderzoek is onderzocht hoe mensen de omgang met een AI-chatbot ervaren vergeleken met een menselijke chat agent in de klantenservice. Het is mogelijk dat mensen in de omgang toch een menselijke medewerker prefereren, omdat zij meer vertrouwen hebben in een medemens en het toegankelijker is om met hen te communiceren. Ten tweede komen antropomorfe chatbots vaker vriendelijker en empathischer over. Desondanks kunnen antropomorfe kenmerken bij een chatbot in sommige gevallen juist averechts werken en mensen ongemakkelijker maken. In het andere geval, wanneer mensen te maken krijgen met een chatbot, hebben ze een voorkeur voor de meest menselijke variant ervan, omdat het duidelijker is voor mensen hoe ze met hen kunnen communiceren.

Tot slot tonen de bevindingen aan dat mensen over het algemeen een voorkeur hebben voor menselijke chat agents. Wanneer ze echter met een AI-chatbot te maken krijgen, hebben ze een voorkeur voor een embodied chatbot met antropomorfe eigenschappen, vergeleken met een disembodied chatbot zonder antropomorfe eigenschappen.

6. Referentielijst

- Adam, M., Wessel, M., & Benlian, A. (2020). AI-based chatbots in customer service and their effects on user compliance. *EM*, 31(2), 427–445. <https://doi.org/10.1007/s12525-020-00414-7>
- 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, 183–189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Ashfaq, M., Jiang, Y., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics And Informatics*, 54, 101473. <https://doi.org/10.1016/j.tele.2020.101473>
- Awasthi, B. (2020, november). *7 Live chat statistics you can't miss*. Software Advice. Geraadpleegd op 13 maart 2024, van <https://www.softwareadvice.com/resources/live-chat-statistics/>
- Brandtzæg, P. B., & Følstad, A. (2017). Why people use chatbots. In *Lecture notes in computer science* (pp. 377–392). https://doi.org/10.1007/978-3-319-70284-1_30
- Chen, S., Li, X., Liu, K., & Wang, X. (2023). Chatbot or human? The impact of online customer service on consumers' purchase intentions. *Psychology & Marketing*, 40(11), 2186–2200. <https://doi.org/10.1002/mar.21862>
- De Cicco, R., Silva, S. C. E., & Alparone, F. R. (2020). Millennials' attitude toward chatbots: an experimental study in a social relationship perspective. *International Journal Of Retail & Distribution Management*, 48(11), 1213–1233. <https://doi.org/10.1108/ijrdm-12-2019-0406>

- Dias, P., & Correia, M. H. (2021). Using Chatbots for Customer Care. In *Advances in marketing, customer relationship management, and e-services book series* (pp. 14–28). <https://doi.org/10.4018/978-1-7998-4772-4.ch002>
- Folk, D. P., Yu, S., & Dunn, E. W. (2022). Is a good bot better than a mediocre human? Chatbots as alternative sources of social connection. *University Of British Columbia*. <https://doi.org/10.31234/osf.io/xunsy>
- Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm. *Human Machine Communication Journal/Human-machine Communication Journal*, 1, 71–86. <https://doi.org/10.30658/hmc.1.5>
- Gümüş, N., & Çark, Ö. (2021). The Effect of Customers' Attitudes Towards Chatbots on their Experience and Behavioural Intention in Turkey. *Interdisciplinary Description Of Complex Systems*, 19(3), 420–436. <https://doi.org/10.7906/indecs.19.3.6>
- Helms, M. M., & Mayo, D. T. (2008). Assessing poor quality service: perceptions of customer service representatives. *Managing Service Quality*, 18(6), 610–622. <https://doi.org/10.1108/09604520810920095>
- Ho, R. C. (2021). Chatbot for Online Customer Service. In *Advances in e-business research series* (pp. 16–31). <https://doi.org/10.4018/978-1-7998-7603-8.ch002>
- Kinna, A. (2005), “The seven deadly skills of customer service”, *The British Journal of Administrative Management*, August/September, pp. 14-15.
- Leggett, K. (2017, juli). *Online self service dominates yet again. Why? Its an effortless way to get to your answers*. Forrester. Geraadpleegd op 13 maart 2024, van https://www.forrester.com/blogs/16-01-28-online_self_service_dominates_yet_again_why_its_an_effortless_way_to_get_to_your_answers/

- Lei, S. I., Shen, H., & Ye, S. (2021).3 *International Journal Of Contemporary Hospitality Management*, 33(11), 3977–3995. <https://doi.org/10.1108/ijchm-12-2020-1399>
- Li, C., & Zhang, J. (2023). Chatbots or me? Consumers' switching between human agents and conversational agents. *Journal Of Retailing And Consumer Services*, 72, 103264. <https://doi.org/10.1016/j.jretconser..103264>
- Liebrecht, C., & Van Hooijdonk, C. (2020). Creating Humanlike Chatbots: What Chatbot Developers Could Learn from Webcare Employees in Adopting a Conversational Human Voice. In *Lecture notes in computer science* (pp. 51–64). https://doi.org/10.1007/978-3-030-39540-7_4
- Morgan, B. (2023, augustus). What impact will AI have on customer service? *Forbes*. Geraadpleegd op 18 maart 2024, van <https://www.forbes.com/sites/blakemorgan/2023/08/16/what-impact-will-ai-have-on-customer-service/?sh=20c1edb56aa6>
- Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal Of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Nguyen, V. T., Phong, L. T., & Khanh, N. T. (2023). The impact of AI chatbots on customer trust: an empirical investigation in the hotel industry. *Consumer Behavior in Tourism And Hospitality*, 18(3), 293–305. <https://doi.org/10.1108/cbth-06-2022-0131>
- Roy, R., & Naidoo, V. (2021). Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *Journal Of Business Research*, 126, 23–34.
- Sands, S., Ferraro, C., Campbell, C., & Tsao, H. (2020). Managing the human–chatbot divide: how service scripts influence service experience. *Journal Of Service Management*, 32(2), 246–264. <https://doi.org/10.1108/josm-06-2019-0203>

- Sanglé-Ferrière, M., & Voyer, B. G. (2019). Friend or foe? Chat as a double-edged sword to assist customers. *Journal Of Service Theory And Practice*, 29(4), 438–461.
<https://doi.org/10.1108/jstp-10-2018-0235>
- Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal Of Management Information Systems*, 37(3), 875–900.
<https://doi.org/10.1080/07421222.2020.1790204>
- Song, M., Xing, X., Duan, Y., Cohen, J., & Mou, J. (2022). Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention. *Journal Of Retailing And Consumer Services*, 66, 102900.
<https://doi.org/10.1016/j.jretconser.2021.102900>
- Svenningsson, N., & Faraon, M. (2019). Artificial Intelligence in Conversational Agents. *AICCC '19: Proceedings Of The 2019 2nd Artificial Intelligence And Cloud Computing Conference*. <https://doi.org/10.1145/3375959.3375973>
- Yang, Q., Steinfeld, A., Rosé, C. P., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings Of The 2020 CHI Conference On Human Factors in Computing Systems*.
<https://doi.org/10.1145/3313831.3376301>
- Yen, C., & Chiang, M. (2020). Trust me, if you can: a study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments. *Behaviour & Information Technology*, 40(11), 1177–1194. <https://doi.org/10.1080/0144929x.2020.1743362>