Gender Stereotypes in Virtual Agents

by

Procheta Nag

B.Sc., (Computer Science and Engineering), Khulna University, 2017

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
School of Interactive Arts and Technology
Faculty of Communication, Art and Technology

© Procheta Nag 2020
SIMON FRASER UNIVERSITY
Fall 2020

Declaration of Committee

Name: **Procheta Nag**

Degree: **Master of Science**

Thesis title: **Gender Stereotypes in Virtual Agents**

Committee: **Chair:** Cheryl Geisler
Professor, Interactive Arts and Technology

Steve DiPaola
Supervisor
Professor, Interactive Arts and Technology

Carman Neustaedter
Committee Member
Professor, Interactive Arts and Technology

Angelica Lim
Examiner
Assistant Professor, Computing Science

Ethics Statement

SFU

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

Visual, behavioural and verbal cues for gender are often used in designing virtual agents to take advantage of their cultural and stereotypical effects on the users. However, recent studies point towards a more gender-balanced view of stereotypical traits and roles in our society. This thesis is intended as an effort towards a progressive and inclusive approach for gender representations in virtual agents. The contributions are two-fold. First, in an iterative design process, representative male, female and androgynous embodied AI agents were created with few differences in their visual attributes. Second, these agents were then used to evaluate the stereotypical assumptions of gendered traits and roles in AI virtual agents. The results showed that, indeed, gender stereotypes are not as effective as previously assumed, and androgynous agents could represent a middle-ground between gendered stereotypes. The thesis findings are presented in the hope to foster discussions in virtual agent research and the frequent stereotypical use of gender representations.

Keywords: Virtual Agent; Gender Stereotype; 3D embodied agent; Quantitative Method

Dedication

This thesis work is dedicated to my parents who have always loved me unconditionally and their upbringing has taught me to work hard for the things I aspire to achieve. This work is also dedicated to my senior supervisor, who has been a constant source of support and encouragement during the challenges of graduate school and life. Thank you for being the role models and sounding boards I have needed.

Acknowledgements

I want to thank my senior supervisor, Professor Steve DiPaola, for his guidance and patience. I have taken more significant strides as a student, as a professional, and as a person since we began to work together. I have learned from you, and I hope to continue learning from you in the future. Thank you for everything.

I also want to thank my committee member, Carman Neustaedter, for his influence and expertise. This research began to take shape in that first research methods class, and it has since grown into something enriching.

Finally, I would like to recognize my lab mates in the ivizlab and in SIAT in general.

Research Acknowledgements

Portions of my research were conducted in collaboration others; here, I make clear the role of each person.

I owe great thanks to my senior supervisor, Professor Steve DiPaola, for his continuous support and guidance in my research work where he acted in an advisory role.

I owe great thanks to Dr. Ozge Nilay Yalcin, a PhD and Post-Doc at that time in our lab. she guided me to integrate my study 3 survey from survey-monkey to Mturk. The interactive Embodied Conversational Agents (ECA) system used to create stimuli of study 1 is our lab's system where she is the main researcher.

Publications

Materials, ideas from this thesis have appeared in the following publication:

Procheta Nag and Özge Nilay Yalçın. 2019. Effect of Gender on Interaction with Embodied Conversational Agents, Women in Machine Learning Workshop, In NeurIPS 2019.

Procheta Nag and Özge Nilay Yalçın. 2020. Gender Stereotypes in Virtual Agents. In Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (IVA '20). Association for Computing Machinery, New York, NY, USA, Article 41, 1–8. DOI:https://doi.org/10.1145/3383652.3423876

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

This work is motivated by what I know through my own experiences. I found that most popular digital voice assistants like Microsoft's Cortana, Apple's Siri (recently added male voice also), Google' Assistant, Amazon's Alexa and other virtual personal assistants have female voices (and generally female names). Additionally, when I was exploring the application of intelligent virtual assistant (AI assistant) or virtual agent for my research work, I noticed a similar pattern. The usage of female agents on certain domains is significantly higher than male agents. At this point, I began to become quite curious to know why there is a gender imbalance and how it can affect human interaction with agents.

In human communication, sometimes women prefer to talk with a female person for a particular topic rather than a man, and a man prefers a male person for talking about a specific topic. This is the natural gender-stereotypical habit of humans. Therefore, it is essential to use a good, embodied character agent on the proper agent-human interaction application. Because people have gender stereotypes (Basow, 1992) and these stereotypes are formed very fast, right after you see someone's face (Jackson, 1992), people tend to assign human-like characteristics to virtual agents, such as gender, personality and emotions - even from a minimal set of behaviours (Reeves, 1998). Therefore, it is clear that the gender of the agent plays a significant role in communication and interaction, such as virtual face to face conversation or indirect conversation, because it will affect people's expectations and how people perceive agents.

According to previous research, gender stereotypes have been found in virtual agent research; however, this trend is changing. Most previous studies used cartoonist character,

avatar, static image of 2D or 3D character, a stick figure of a character with a traditional chat-bot system to validate the gender stereotype assumptions. But they do not have any embodied activities rather only their static physical appearance (a still image). On the other hand, some studies used dynamic embodied agents but have limited verbal and non-verbal behaviours. Some studies do not use voice for back and forth conversations with the user. Instead, these studies used text input or multiple choice input from the user and responded with voice and used recorded voice for agent speech rather than synthesize voice with real-time lip synchronization. Moreover, most of the studies have adopted only the binary gendered character to validate the gender stereotype assumption.

Against this backdrop, this thesis presents a series of studies to evaluate the stereotypical beliefs of gender by systematically manipulating the visual cues of human-like embodied virtual agents. The approach to achieve the intended research goal can be summed as follows: first, I have created representative human-like male, female and androgynous agent characters with a few differences in their visual attributes other than gender as possible through an iterative design process. Then these agents are used to evaluate the stereotypical assumptions of gendered traits and roles in virtual agents. The study results showed that gender stereotypes are not as effective as previously assumed, and androgynous agents could represent a middle-ground between gendered stereotypes. Unlike previous work, this research evaluated the gender perception of my designed agents (created/validated specifically for this task) before evaluating the gender-based stereotypes, used realistic-looking 3D characters instead of still, drawn or cartoon-like (or even borrow from the commercial system) representations and included androgynous agents in this virtual agent gender stereotype evaluation study.

1.1 Background

Historically, women are often seen in the assistant roles, reception and help desk, and hospitality roles. In more recent history, the voice user interface evolution, like phone call help service and digital announcements, has been widely used. The telephone operators have traditionally been female, making people accustomed to getting assistance from a

woman's voice. Using female-gendered personalities for voice user interfaces is not new. Navigation systems are, for example, also typically equipped with female-gendered voices such as Google Maps. A recent UNESCO report (West et al., 2019) reveals that the most popular voice-based conversational agents reinforce the gender bias in social context (Specia, 2019). Whereas the conversational user interface of Google Assistant does not have a feminist name, but its voice is female. There can be something biological and socially based behind the sources, preferring female gender representation for assistance work (Griggs, 2011).

With recent developments in computer graphics and animation, many conversational agents under development today are fully embodied (Cassell et al., 2000) with a human-like interface and endowed with a wide range of human-like behaviours. Hence, plenty of talking heads and virtual embodied agents are currently interacting with users through their voice and virtual bodies, physical gestures and postures, especially on the Internet. There has been a rising prominence of the embodied conversational agent (ECA) research and applications in the social context, conversational context and interface design.

Virtual agents (VAs) are artificial agents that use natural input and output methods, allowing for an effortless interaction similar to human-to-human dialogue (Massaro et al., 1999). I will use this as an umbrella term for all forms of these types of agents. These agents can be in the form of chatbots that use text-only interaction, Voice-based agents that use speech as input and output modalities, or Embodied Conversational Agents (ECAs) multi-modal interaction with a virtual representation, such as an avatar. The virtual physical form or embodiment of ECAs is commonly represented by three-dimensional human-like characters with human-like verbal and non-verbal behaviours to enable face-to-face conversation with users, usually on a digital screen, providing a powerful means of online communication. Some advanced systems render their 3D agents in Virtual Reality (VR) or Augmented Reality (AR) displays. However, their design and behaviour are mostly projected as gendered characters. ECA's can usually understand both spoken and written inputs but are mainly designed for spoken input to respond with voice (computer-generated voice or human voice). The non-verbal information like facial expressions, body postures, gestures, or voice tone of both the agent and the human user is utilized to compensate for the conversation dialogue

system. Typically these agents' actions and outputs are directed by complex AI systems, natural language processing, self-learning, and machine learning to learn as they go and provide a personalized, conversational experience to support a wide range of user queries and other scenarios such as being coaches, teachers or narrative players.

Embodied conversational agents (ECAs) are unique (Venkatesh et al., 2018) from regular interface agents, chatbots and popular voice assistants because the regular interface agents are usually represented by cartoonist or two-dimensional characters with some simple mouth or hand gestures during communication with users. Though the agents communicate with users through speech, the virtual physical form is often a human or sometimes a non-human projection (Forlizzi et al., 2007).

Whereas, Voice Assistants respond to voice orders and give the user data about their inquiry through voice, but they do not project a physical form. Though they can receive both speech and text inputs to generate voice response, the primary application purpose is for simple informational back and forth voice interaction with the user. Currently, voice assistants can answer queries, perform entertainment activities like playing music, act as a smart home assistant, and call a device contact list. From this revolutionary decade of voice assistants, the most popular voice assistants of Apple's "Siri", Amazon's "Alexa", Microsoft's "Cortana" and Google's "Google assistant" offer service with female voice output (West et al., 2019). On the other hand, chatbots interact with users primarily through written text rather than spoken language. The representation of a chatbot can vary. For example, the chatbot of IBM's Watson assistant (IBM, 2018) does not project any visual character. When a physical form is projected, it is ordinarily static – often a still image of a human face or sometimes a non-human image, such as a cartoon character. For example, though "Mitsuku" is a conversational chatbot, it projects an animated portrait of a teenage female cartoon character in the chatting page (PANDORABOTS, 2013). It should be noted that the agent terms I use here, while generally accurate, have at times been giving broader meanings, mainly because of the new and emergent nature of the area.

Thus, ECAs are distinct from all other agents mentioned above because their output is a multi-modal integrated humanoid depiction that is human-like figures displaying human-

like synchronized gestures, behaviours and lip-synced speech, not just non embodied voice or written text. In other words, the advanced empathetic expression, verbal and non-verbal behaviours with speech appear to emanate from someone like a human figure a user can see and interact with, usually on a digital screen or in a virtual or augmented reality environment, have made them unique. The nonverbal channels are essential not only for conveying information but also for regulating the conversation flow. The nonverbal channel is particularly essential for social dialogue to replicate the human-like conversation. Though researchers have analyzed the impact of gender stereotypes in voice user interfaces, this impact has yet to be analyzed for embodied agents in all aspects as it holds more stereotypical attributes than just only chatbot or voice assistant. According to Reeves, as humans tend to anthropomorphize computer agents, they also assign human-like characteristics to these non-biological beings such as age, nationality, race and gender (Reeves, 1998) to make embodied forms. The stereotypes related to these characteristics are also attributed to these agents due to this unconscious assignment.

One of the most prominent of these stereotypes is gender-related stereotypes. Multiple studies suggest that gender characteristics affect the user experience depending on the role of the agent (Baylor and Kim, 2004; Brahnam and De Angeli, 2012). The social-role theory suggests the gender stereotypes originate from the observation of different genders in their social functions (Eagly and Wood, 2012). Female stereotypes have been dominated by affectionate, warm, expressive communication behaviour, which are reflected in social roles as assistants, caretakers and companions. However, male stereotypes are dominated in strong, bold and dominating behaviours patterns such as mentors, motivators and exercise coaches. Mirroring these stereotypes, we see a surge in predominantly female gender cues in both academic and commercial assistant agents (e.g., Microsoft Cortana, Apple Siri, Amazon Alexa, Google Assistant) (Bresge, 2019; LaFrance, 2016).

These different behavioural patterns for different gender develop stereotype traits among people. These traits are mostly categorized as "communion" and "agency," where "competence" is also considered by some as a distinct trait (Bakan, 1966; Sczesny et al., 2018). Communion trait includes qualities such as compassionate, warm and expressive; and is

considered to be a stereotypical trait of the female gender. Agency includes qualities such as ambitious, assertive and competitive, which is regarded as the dominant male stereotypical trait. Competence (e.g., intelligent, innovative) has been conceptualized differently based on the domain and suggested to be correlated with both agency, warmth or even morality (Abele et al., 2016). These traits are universal across almost all the cultures (Williams and Best, 1990). Traditionally, in many cultures, any man/woman not conforming to these gender roles would be considered less of a man/woman and shamed.

1.2 Thesis Problems

This thesis investigates the design and evaluation of embodied agents around gender stereotypes. Embodied virtual agents (ECAs) are computer-generated animated anthropomorphic interface agents which can engage a user in real-time, multi-modal dialogue, using speech, lip synchronization, gesture, facial expression, gaze, posture, intonation, and other verbal and nonverbal behaviours to exhibit the experience of face-to-face conversational interaction like human interaction (Cassell et al., 2000; Huang, 2017) (see Section 1.1 to review about ECAs). The overarching research problem of this thesis is: we do not know whether the gendered appearances of agents alone affects the perception of gender stereotypical traits or the preference over certain stereotypical roles. To more easily address this overarching research problem, I have broken my main problem into three sub-problems:

Research Problem 1: When only varying with binary gender-related appearance of embodied virtual agents (ECAs), how does the experience of the user change?

The findings in most virtual agent research and studies show extensive usage of female agents in commercial systems. This overuse of a specific gender in ECAs appears to give a warning sign of the gender stereotypes issue (Khan and De Angeli, 2009; Pelachaud, 2015; Ring et al., 2016; Schulman and Bickmore, 2009; Wang et al., 2015; Silvervarg and Jönsson, 2011; Zhou et al., 2018; Bickmore et al., 2018, 2013a; Trinh et al., 2018; Bickmore et al., 2010b; Utami et al., 2017; Bickmore and Cassell, 2005). On top of that, l (Payne et al., 2013; Louwerse et al., 2005; Gulz et al., 2007; Koda and Takeda, 2018). While many

previous studies show that users prefer female over male agents, in the study of Andre and Baylor (Baylor et al., 2003a; André and Pelachaud, 2010) audiences preferred a male tutor when it comes to technical or engineering subjects over a female tutor. However, the results from the more recent scientific literature on virtual agents show a changing trend on gender stereotypes in recent years (Eagly et al., 2020). Using realistic 3D ECAs, this thesis investigates these current gender stereotype trends by varying gender representation in user studies.

Research Problem 2: Regarding moving past binary distinctions of male and female, how does the more fluid notion of gender (or even gender neutral) affect a user's experience with ECAs?

The modern concepts of gender fluidity and androgyny have not been well studied in the virtual agent research community. The recent increase in awareness of non-binary gender categorization and movements makes it essential to approach androgynous agent character and male and female. Very little research has been done on providing guidelines to create an androgynous character for a virtual agent. Therefore, researchers need to investigate this area starting with how to represent the androgynous character's visual appearance for an embodied agent.

Research Problem 3: How do the gender-related appearance of male, female and androgynous agents impact the user perception in terms of gender stereotypes traits?

The dominance of using female cues in agents in assistant roles received much criticism (West et al., 2019) as these agents are thought to enforce social bias and do more harm than help in closing the gap on gender inequality (Specia, 2019). Moreover, recent research suggests that these gender roles are fluid, and the differences between gender stereotypes are not as biased in our society. These stereotypes tend to reflect on the underlying traits assigned to different genders, which are a direct result of observing different genders in their social roles (Eagly and Wood, 2012). These traits are mostly categorized as "communion" and "agency," where "competence" is also considered by some as a distinct trait (Bakan, 1966; Sczesny et al., 2018). A comprehensive meta-analysis of Eagly et al. has shown a

dramatic change in stereotypical traits and functions in the last several decades (Eagly et al., 2020). Therefore, it is important to reevaluate the user experience towards embodied agents in terms of gender stereotypes traits and roles (see Section 2.3 and 1.1 to review).

These above factors indicate the need for a revised look at gender cues' effect in virtual agents. The old assumptions and research need to be revisited in light of new evidence on the critical gap in gender stereotypes. Moreover, gender appearance on perceived stereotypical traits and gender roles has not been previously studied in virtual agent research.

1.3 Thesis Goals

My primary goal with this thesis is to present a series of studies to evaluate gender's stereotypical gender assumptions by systematically manipulating embodied virtual agents' visual cues. This thesis includes the research of human-like agents (not including the non-human/robot/animals or other virtual characters). This research investigates to validate the gender stereotypes assumption, iterative process for creating examples of gendered agents with minimal differences, and the evaluation of these characters in terms of gender stereotypes and roles. To address this overarching goal, I seek to address the following objectives, which are aligned with the thesis mentioned above problems:

Objective 1: Examine the effect of gender during a user's conversation with the embodied conversational agent.

I will examine user change experience when presented with gender distinctions of embodied conversational agents (ECAs). Chapter 3 presents the study to validate the gender stereotypes trend proven by numerous previous research. Though some previous research has shown gender stereotypes studying virtual agents, very few have used the embodied agents that can perform face to face conversation in real-time through verbal and non-verbal behaviours to validate the trend. In my study, I have presented the embodied conversational agents full of gestures and facial expressions and can continue the conversation with interaction partners through voice.

Objective 2: Examine the user experience on perceiving realistic 3D ECA characters, where all other factors are kept as uniform as possible, to measure for differences in ECA visual gender cues only.

The design requirements of virtual agent characters emerged from study 1 and related literature described in chapter 2. The significant drawbacks of study 1 were - not including an androgynous/gender-neutral agent to evaluate the gender stereotype assumption and the difference between the agents' visual appearance indicates the importance of balanced visual design, such as an imbalance between the skin colour, hair colour, and clothing colour scheme. From the literature research, I have seen that most previous studies used cartoonist character, avatar, static image of 2D or 3D character, a stick figure of a character with a traditional chat-bot system to validate the gender stereotype assumptions and adopted only binary gendered characters mostly to validate the gender stereotype assumption. To alleviate this research gaps of 1) very different styles of ECA characters and 2) too many confounding variables other than gender appearance, in chapter 4, I have created binary and non-binary highly configured (e.g. high-end 3d game character level) characters of agents by investigating the gendered 3D model's essential features from the user perspective and improved them through design.

Objective 3: Examine the effect of the agents' gender-related appearance: male, female and androgynous, on user perception in terms of gender stereotype traits and roles.

Chapter 5 presents the third study that examines the general population's point of view (from our more open study participants) on changing gender stereotypes, changing gender roles, and a dynamic understanding of the notion of gender. The goal is to correlate these modern views of gender with the current research. The three main categories of stereotypical traits (i.e., agency, communion and competence) and examples of roles will be examined, supported by the recent statistics on the labour force (Statistics., 2019; STATISTICS, 2020) and changing gender stereotypes perspective research by Eagly et al. (Eagly et al., 2020). Study III is intended as an effort towards a progressive and inclusive approach for gender representations in virtual agents. However, Eagly et al. (Eagly et al., 2020) have done a

meta-study; they did not run an actual analysis with a virtual agent to support their claim. I have employed their findings in virtual agent research and conducted a comprehensive study using the high-end realistic agents I created in chapter 4 to evaluate their claim.

1.4 Thesis Organization

In Chapter 2, I provide a literature review. I discuss how gender stereotypes are happening in society and human life. I provide a detailed literature review of the existing research application of virtual agents in different domains. And lastly, I describe and analyze the design of the virtual agent characters of these existing research.

In Chapter 3, I present Study 1 (Gender preferences for ECAs) which employs a quantitative research method to examine the effect of gender during the interaction with the embodied conversational agent. The study uses a survey method to gather data from participants to see if gender stereotypes exist in the agent's preference in an assistant role. Eventually, I discussed the findings and lessons that I learned from the study to improve in the latter study.

In Chapter 4, I present Study 2 (Character design and evaluation) that exhibits the design of agent characters, a male, female and an androgynous label. First, I describe the design requirements which emerged from the lessons I learned from Chapter 3 and related literature mentioned in Chapter 2. Realizing the issues with other studies that often use VA characters that are just available or random or where female and male characters differ by more than just the gender controlling variable, I more systemically created the characters and evaluated and validated those designs through a survey study.

In Chapter 5, I present Study 3 (Evaluation of stereotype traits) that examines gender stereotypical traits and roles based on the agents' gendered appearance. I adopted a quantitative research method for the study and used a male, female, androgynous character created in Chapter 4. At first, I explain how the domain and roles are chosen for the agent, then describe the study's design and evaluation, and finally discuss the study's findings. Then I discuss the results and identify the design implications of embodied agents for fu-

ture systems. This involves examining the positives (what worked out well) and negatives (what did not work out well) considering the limitations of my design and study.

In Chapter 6, I conclude my thesis by revisiting my research goals and reflecting on how I achieved each one. I also list my research contributions and suggested areas for future work.

Chapter 2

Related Works

The subsequent sections include a brief review of the related work on gender stereotypes in virtual agent research.

2.1   Gender Stereotypes

A person's gender refers to characteristics related to male, female, or some combination thereof. Although the words "sex" and "gender" are often used together for the same purpose, they have different meanings. The concept of gender is fluid and varies from society to society. Judith Butler clarified the distinction between sex as a natural given category and gender as an acquired cultural - social category (Butler, 2011). Sex is determined by a person's sexual reproduction organ, whereas a person's gender is influenced by environment, society and biological factors. Gender is how a person psychologically and emotionally decides to identify themselves. A person who does not conform to any traditional gender is called a gender-nonconforming or non-binary person.

Stereotypes are a cognitive structure containing knowledge, expectations, and beliefs about social category (Hamilton and Trolier, 1986). It is a generalized and simplified fixed idea of a specific group of people. Stereotypes can be explicit or implicit, individual or cultural. A gender stereotype is a set of predetermined views on how a person should behave according to their assigned gender. Gender stereotypes have existed in society since the earliest human communities. This set of ideas are implanted in a person's mind by society and culture, and it is activated as soon as the person sees another human being (Jackson, 1992). These stereotypes often disregard the existence of any gender other than men and

women. Gender stereotypes can be categorized as descriptive stereotypes regarding what characteristics men and women possess and prescriptive stereotypes as in a set of characteristics describing how men and women should be (Burgess and Borgida, 1999). People's insights and their reflection and behaviours towards others are thus dictated by their cognitive representations of gender and their normative beliefs about what is appropriate and inappropriate for other people to do given their gender (Heilman, 2001).

Traditionally minded stereotypical traits attributed to men would be aggressiveness, dominance, sexuality, independence, whereas women are stereotyped to perform as kind, affectionate, submitting and subservient, soft-spoken, communicative (Cameron, 2001). These traits are universal across almost all the cultures (Williams and Best, 1990). Traditionally, in many cultures, any man/woman not conforming to these gender roles would be considered less of a man/woman and shamed.

2.1.1 Theoretical Aspect of Gender Stereotypes

Social Agency Theory

Social agency theory posits that using verbal and visual social cues in computer-based environments can foster a partnership by encouraging the learners to consider their interaction with the computer to be similar to what they would expect from a human to human conversation (Argyle, 1988). The environment relies on expressed gestures such as a standard accented voice or non-verbalized gestures such as facial expressions to stimulate the learner to engage in the situation as though they are involved in a human to human interaction. This theory posits that gender stereotypes are planted in the human brain from a very young age. A child's behaviour is typically controlled by stereotypes of their gender that have been imparted upon them by the society (Burgess and Borgida, 1999).

Social Role Theory

Eagly's social role theory argues that widely shared gender stereotypes develop from the gender division of labour that characterizes a society (Eagly and Karau, 2002). This gender division of labour hands men and women different work skills. The expected behaviour of men and women is projected by social scripts that manipulate relational communica-

tion between people of the same and different genders. These social scripts are related to stereotypical personality traits for all genders. Men are supposed to be more assertive and proficient in a social setting, whereas women are supposed to be more docile. Men are understood to project more agency-oriented independence and success, while women are advised to exhibit more communal oriented activities with interpersonal support.

Similarity Attraction Theory

A common belief in humankind is that people are induced by others, want to be similar to others in different positive aspects. Researches on opposite-sex attraction reveal that man valued appearance and attractiveness more than women. Women, on the other hand, appreciated similarities and interests more than men (Stroebe et al., 1971). Also, it has been concluded that, under certain circumstances, people are less influenced by similarities, especially when similarities are combined with negative traits such as "unattractiveness," "mental illness," or "unpleasant" behaviour in general (Feingold, 1994). In most cases, people are more likely to be influenced by social models interested in them or similar to how they would like to be.

Attractiveness Stereotype Theory

A person's first impression of someone is often fashioned by the other person's appearance and face. A traditionally "attractive" person would receive more positive and warm reactions from society rather than a conventional "unattractive" person (Chaiken, 1979). Many facial features are thought to work in conjunction with making a person attractive, such as averageness, symmetry, and lack of blemishes. The assessment of attractiveness also depends on gender. Traditionally, men are believed to prefer prominent eyes and cheekbones in women, while females tend to take in jaw width and lower-face proportions to weigh attractiveness in males. In addition to that, men consider physical attractiveness more important when judging women than women do when evaluating men (Feingold, 1994). Attractiveness is also strongly associated with positive traits, especially with interpersonal competence, social appeal, and adjustment (Eagly and Karau, 2002). Although there is less research on the assessment of "unattractiveness," it is inevitable that negative human traits are associated

with "unattractiveness." "Unattractive" people experience more hostility from society than their "attractive" counterparts, especially women. Women are always expected to preserve a youthful look despite their age, hence the very significant makeup and plastic surgery industries. Women are judged by their attractiveness to a much greater extent than men (Jackson, 1992).

2.1.2 Why a Virtual Agent's Gender Matters?

In recent years, the development of computer technology and advanced AI software has brought a new flock of AI-powered systems. These VA based bots act as the user's friendly virtual assistants. Embodied Conversational Agents (ECA) are an improved and refined version of these VA based bots. These agents are computer-generated animated interfaces that can interact with users in real-time using verbal and nonverbal gestures such as speech, facial expression, posture, intonation etc. They differ from regular chatbots and voice agents because they are more human-like three-dimensional characters represented as ECAs.

In contrast, chatbots and voice assistants are characterized by two-dimensional projections or do not have any physical projection at all (Venkatesh et al., 2018). This three-dimensional physical appearance creates a human-like persona for the agents, which helps them interact with users like they would another regular human being. Through their form, clothing, facial expression, and gestures, embodied agents extract the rich communication channels found in human social interaction into the interface, often reviewed to be more engaging to the user (Moreno et al., 2002). Thus, they open the door for social gender stereotypes to come into play.

According to Nass (Nass et al., 1997), "computer personalities can be easily created using a minimal set of cues. People will respond to these personalities in the same way they would react to similar human personalities". Embodied Conversational Agents stimulates social interaction and communication by their social (character and voice) and visual cues (physical appearance). These human-like personifications mostly bring the gender aspect to ECAs and HCIs, prompting these virtual bodies to carry stereotypical attributions and that users respond differently to ECAs (Brahnam and De Angeli, 2012).

Physical Appearance

Physical appearance serves as a trigger when determining a person's gender. Physical appearances of agents instigate a person to determine the expected behaviour from agents. It also influences the interpretation of actions and intentions, and thus appearance plays a significant role in the outcomes of social actions (Guadagno et al., 2011b). A person's "attractiveness" is also defined by their physical appearance. Interpretation of physical appearance also depends on the context (Lankes et al., 2007). This traditional concept of physical appearance and attractiveness encourages people to prefer attractive virtual agents of the opposite sex. According to Gulz (Gulz et al., 2007) the physical appearance of a synthetic character can be categorized into:

1. Static Visual Characteristics:

    - degrees of anthropomorphism and degrees of visual naturalism or realism.

    - basic physical properties such as body-type, face properties, skin, hair and haircut, clothes and attributes – characteristics that, furthermore, can make up representations of age, gender and ethnicity.

2. Dynamic Visual Characteristics:

    - facial expressions, body and hand gestures, postures, movements, etc.

Attractive agents are considered more persuasive and influential in sales management than their "unattractive" counterparts. Consumers maintain a more positive evaluation of embodied conversational agents rated more attractive than they did with those placed less attractive (Khan and De Angeli, 2009).

Voice

Voice is a very prominent stereotypical attribute of gender. Males are often attributed with a deep husky voice, whereas women are considered high pitched ones. It's a well-established phenomenon that the human brain is developed to like female voices, and It's much easier to find a female voice that everyone wants than a male voice that everyone likes(Nass et al.,

1997). People recognize the gender of a machine when the singer has only the slightest hint of gender. Female voice agents are perceived as much warmer than their counterparts (Vala et al., 2011). Clients appraise the agents with a human voice more compared to machine-synthesized voice (Atkinson et al., 2005). In a CASA study examining whether gender stereotypes extend to machines exhibiting gender cues, male-voiced computers were rated by users to be more proficient than female-voiced computers (Nass et al., 1997).

Context

People's assumptions about the gender of virtual agents can be activated even by a minimal set of behaviours. People tend to prefer agents that conform to the existing gender stereotype and use their current communication and social skills while interacting with artificial agents (Forlizzi et al., 2007). The agents are embodied with natural communication capabilities to create a familiar environment to interact with. These gender embodied agents pursued users in following the advice given by male characters when the subject was masculine, such as sports, and by female characters when the topic was feminine, such as cosmetics and fashion (Gulz et al., 2007). The gender of an educational agent also influences students' expectations and perceptions of the agent, and these expectations impact their learning (Baylor, 2005).

2.2 Agent's physical appearance and context

Social psychology research indicates that humans are more persuaded by a person from a member of their in-group; work with anthropomorphic agents has also shown that observers tend to be more influenced by an agent of the same gender (Baylor, 2005; Kim et al., 2006; Kim, 2016). Zimmerman's study (Zimmerman et al., 2005) compared the agents found in research and industry and his own created designs to explore whether these embodied agents produce the same effects I observed in the first study. The studies' main findings include that people prefer agent forms that conform to gender stereotypes for different tasks and roles, and men prefer embodied agents more than women do. One study found that men and women did not have a gender preference in selecting an embodied agent with a human form (Cowell and Stanney, 2003). Another found that people did not perceive a difference

17

in the intelligence in male and female looking human agents (Koda and Maes, 1996). This research indicates a clear link between the visual appearance of the agent and its application in research.

2.2.1 Person perception of visual appearance

There is a possibility that agents can be represented more naturally to other physical presentations if researchers know how an ordinary individual perceives and recognize different appearance factors. In the present time, agents' external physical/ visual appearance is mostly constructed by designers. User preferences are mostly limited to letting users choose the agent's personalities from a fixed set of lists. Therefore, it is clear that the agent's embodiment relies on human beings to emerge the effect of the agent's physical appearance upon the user. Thus, an individual's perception model will influence agents' social activity and personality they will possess. However, it is not possible to presume all the circumstances and consequences in advance during the design period until the agent user interaction happens.

Therefore, it is necessary to consider several issues while designing a virtual agent and developing the physical form. According to Baylor (Baylor, 2009), when we design human-like avatars and agents to represent ourselves through virtual worlds and gaming environments, their appearance has a considerable impact on our virtual and real-world behaviour. By adding face and embodiment, the social relationship between user and computer becomes even more explicit: clothing, hairstyle, facial expressions, age, gender, etc. displayed by agents bring the rich and complex world of social interactions into the interface (Brahnam and De Angeli, 2012). Other necessary research has demonstrated that people interact with computers as if they were social actors, even when the computer does not take an anthropomorphic or embodied form (Walker et al., 1994). The paper investigated whether a gender-stereotypical pattern of effective evaluations could be elicited when degrees of femininity and masculinity are manipulated visually instead of via voice cues to create gender neutral appearance.

2.2.2 Visual design of Embodied Agent

Very few research and studies have focused on designing gender less or gender-neutral characters such as androgynous characters for the virtual agent to study the general context and roles traditionally presented by male or female agent characters. However, from these previous research, it can be seen that they have tried to create some androgynous characters breaking the traditional pattern. The following table 2.1 summarizes the visual features of male, female, androgynous character, which researchers have followed to create the character for their study.

Image	Male	Female	Androgynous
see figure 2.1 (Gulz et al., 2007)	broader, angular and more pronounced head shapes; broader shoulders, a distinct Adam's apple, pronounced, dark eyebrows, neatly done hair	baby- face scheme (rounded head shapes, bigger eyes, smaller nose, narrower shoulders); long (coloured) hair and makeup, fuller lips	weak male: rounder and less pronounced shapes of head, jaw and nose; narrower shoulders; slightly red lips helpless female: broader head, a more angular and prominent jaw, a high for-head; paler colours as to eyes, mouth and hair
see fugure 2.4a (Silvervarg et al., 2012)	more angular, broad, pronounced shapes (head, eyes, nose, mouth, etc.), straight eyebrows, short hair, and broader shoulders	rounded shapes, big eyes, curved look browses, small nose and mouth (the babyface scheme) as well as pronounced eyelashes, long hair, and narrow shoulders	manipulated to be somewhere between the female and male attributes (e.g. shapes, eyebrows, and mouth). Especially the length of the hair was iterated several times
see figure 2.3 (Niculescu et al., 2009)	darker skin colour, moustache and/or beard and short hair; one of the heads is blond, the other two are brunettes	earrings, have narrow eyebrows, a lighter skin colour and a feminine hairstyle; each head has a different hair colour	young male or a female with slight masculine traits

Table 2.1: Features of the visual appearance of male, female and gender-neutral agents from previous research.

Figure 2.1: Agent's gender related appearance of female (top left), weak female (top right), male (bottom right), weak male (bottom left) character (Gulz et al., 2007).

Figure 2.2: Agent's gender related appearance female, androgynous, male (from left to right) character (Silvervarg et al., 2012).

From the table, 2.1 and figures, some common features to provide the characters specific gender-related appearance can be summarized as-

- Male character: Prominent colour scheme reflects the gender-related appearance of male and combined coloured lips with an overall paler colour scheme to weaken the masculine visualization of character.

- Female character: Rounder face structure with slight makeup is typical to reflect the feminine attributes, and large eyes with fuller lips make it more distinct to stand out as a female character.

Figure 2.3: Agent's gender related appearance of male (top three), female (bottom three) and androgynous (right one) character (Niculescu et al., 2009).

- Androgynous/gender-neutral character: Some distinct features from masculine and feminine attributes are combined to create a gender-neutral representation of the human-like character.

Research shows that the effect of physical appearance changes as indicated by the unique circumstance (Lankes et al., 2007). When embodied conversational agents are presented as gendered human-like characters, users more frequently followed the male characters' guidance when conversational topics were manly, such as sports. Still, they followed female characters when the topic was feminine, such as cosmetics and fashion (Gulz et al., 2007). In a study with college students, Baylor (Kim et al., 2006) found the gender of a pedagogical agent influencing student expectations and perceptions of the agent. Morton (Morton et al., 2004) also report users applying gender stereotypes to pedagogical agents. They discovered that student's sexual orientation desires affected their learning.

2.2.3 Agent's Role and Visual Appearance

A great deal of research has been done on embodied conversational agents in various domains, roles and appearances. However, very few have focused on the stereotypical perspective of the gender-related appearance of ECAs and the effect that appears on the application domain and roles. Many researchers believe that conversational agents offer users significant benefits, including higher levels of engagement and providing users with human-like interfaces that require little or no training to use (Foster, 2007; Dehn and Van Mulken,

2000; André and Pelachaud, 2010; Isbister and Doyle, 2002; Maya et al., 2004; Pelachaud, 2005). People are supposed to interact with these interfaces naturally, just as they would with another person. Virtual bodies afford nonverbal cues that have the potential of enhancing sociality (via facial expression recognition and simulation), and by utilizing the social stereotypes and other "cognitive short-hands" people employ in their everyday encounters with other humans (Cassell, 2001; Laurel and Mountford, 1990; Morton et al., 2004; Zhou et al., 2018).

Many experimental studies have been performed on the application of virtual agents in real life scenarios such as family health coach (Wang et al., 2015), patient's health advisor (Bickmore et al., 2018; Trinh et al., 2018; Zhou et al., 2018; Bickmore et al., 2013a; Schulman et al., 2011), empathetic personal companion (Utami et al., 2017), motivational counselor (Schulman and Bickmore, 2009; Baylor, 2009), depression therapist (Ring et al., 2016), study tutor (Silvervarg and Jönsson, 2011), conversational partner (Khan and De Angeli, 2009; Koda and Takeda, 2018), real state expert (Cassell, 2001), career mentor and police officer (Schwartz, 2020). Surprisingly researchers have used female character for all these roles.

Some of the studies have evaluated the use of male or female-looking agents in educational application as tutor (Baylor and Kim, 2003; Gulz et al., 2007; Schroeder and Adesope, 2015; André and Pelachaud, 2010), motivational speaker (Baylor, 2005, 2009) and in entertainment application as chatting partner (Castillo et al., 2018; Koda and Maes, 1996), presenter (Louwerse et al., 2005). These studies have shown interesting gender related findings from their study. Baylor showed in his paper (Baylor, 2009) that even though students preferred cool, attractive, young female agent, at times they also liked the teaching of unattractive, uncool male agent.

Though, very few studies have focused on gender neutral or androgynous agent character. Silver (Silvervarg et al., 2012) presented a visually androgynous ECA character as a teachable agent in an educational math game and compared it with a female and male agent. The results showed that visually androgynous agent was less abused than the female but more than the male agent. Regardless of gender, students showed a significant prefer-

ence for the androgynous agent over the gendered (female and male) agents. These results suggest that androgyny may be a way to keep both genders represented.

Several studies have shown the potential use of conversational agents in various domains and roles such as health, education, assistance, entertainment and so on. In the following tables 2.2, 2.3 and 2.4, I have presented an overview of the agents appearance regarding the roles they are given. The appearance includes the gender representation, the embodiment of how the physical appearance has been presented either head-shot or a full body or static illustration.

Educational Domain

In the educational domain, embodied virtual agents perform various roles in academia such as a tutor, academic instructor, academic expert, motivator/motivational speaker, career guide, presenter, lecturer. Andre, in this study (André and Pelachaud, 2010) chose young agents to exhibit friendly voice and youthful appearance of a motivator because the target audience of the study was the young generation. Though it's being reported in previous studies that users prefer female over male, in the study of Andre and Baylor's (Baylor et al., 2003a; André and Pelachaud, 2010) audience preferred the male tutor when it comes to technical or engineering subjects over female tutor. Table 2.2 and figure 2.4 describes the representation and activities of the virtual agents that previous studies have used for education domain research application.

Author	Gender	Embodiment	Role	Task	Age
(Silvervarg and Jönsson, 2011)	female	headshot, cartoonist, still image (figure 2.4b)	tutor	tutoring teenage students educational math game	young
(Massaro, 2004)	male	headshot, 2D, cartoonist, still image	tutor	tutoring students	adult
(Schroeder and Adesope, 2015)	male, female	full body, 2D, cartoonist (figure 2.4d)	instructor	provided the instructional material to the learner	adult

Continue to next page...

Author	Gender	Embodiment	Role	Task	Age
(Silvervarg, 2016)	andro - gynous	full body, cartoonist, interactive (figure 2.4a)	educa- tional software	students interacted with educational software agents during four hour-long lessons	young
(André and Pelachaud, 2010)	female	upper body, 2D, cartoonist (figure 2.4e)	engineer- ing motivator	delivered the persuasive message about engineering	young
(Baylor et al., 2003a)	male (2), fe- male(2)	head shot, 3D	technol- ogy trainer	technology tutoring session	adult
(Baylor, 2005)	male, female	head shot, 2D, cartoonist (figure 2.4c, 2.4g, 2.4h)	academic expert, motivator	presentation on coping with college life	adult
(Baylor, 2009)	female, male	portrait, 2D, cartoonist	educative motiva- tional speaker	motivational speech	adult
(Wik and Hjalmars- son, 2009)	male	head shot, 2D, cartoonist (figure 2.4f)	learning tutor	teach computer language	adult
(Gulz et al., 2007)	male, feminine male, female, mascu- line female	upper body, 3D	medical tutor	lecture (flash movie) young students	adult
(Baylor and Ebbers, 2003)	male	head shot, 2D	motivator, expert, mentor	teach instructional planning for educational technology course	adult
(Silvervarg et al., 2013)	male, female, androgy- nous	head shot, cartoonist, illustration	teacher	multiple choice chatting conversations with students about math game	young

Table 2.2: Table shows details of the agents' appearance and different roles in the education domain.

(a) (Silvervarg, 2016)

(b) (Silvervarg and Jönsson, 2011)
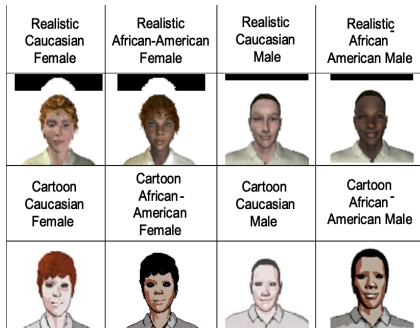
(c) (Baylor, 2005)
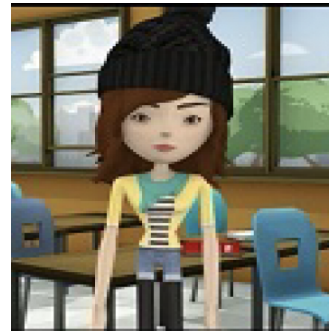
(d) (Schroeder and Adesope, 2015)

(e) (André and Pelachaud, 2010)

(f) (Wik and Hjalmarsson, 2009)
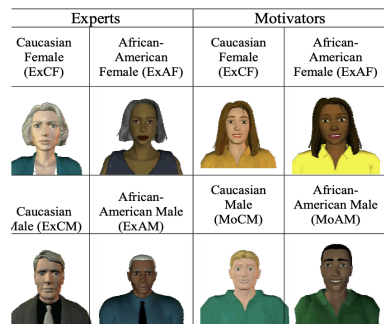
(g) (Baylor, 2005)

(h) (Baylor, 2005)

Figure 2.4: Agent characters used in previous research in educational domain.

Health Domain

In the health domain, embodied virtual agents are presented performing different roles such as health counsellor, consultant, advisor, health guide, health coach, nutritionist, and so on. Table 2.3 and figure 5.7 describes the representation and activities of the virtual agents that previous studies have used for health domain research application.

Professional and Entertainment Domain

Virtual agents are presented in the embodied or disembodied form in the entertainment domain, performing different roles such as personal assistant, conversation partner, and chatting partner. Agents are also performing very professional functions like salespeople, real estate agents, museum guides and event presenters. It is essential to mention that Louwerse (Louwerse et al., 2005) conducted a study comparing the agent's cartoonist character and realistic on the gender and realism created, and the findings showed that people were more approachable towards the more realistic character. Table 2.4 and figure 2.6 describes the representation and activities of the virtual agents that previous studies have used for professional and entertainment domain research application.

Author	Gender	Embodiment	Role	Task	Age
(Ring et al., 2016)	female	upper body, 3D, cartoonist (figure 2.5a)	coun-selling	depression counselling	adult
(Schulman and Bickmore, 2009)	female	full-body, cartoonist, 2D (figure 2.5b)	coun-selling	motivational counselling	adult
(Wang et al., 2015)	female	upper body, 2D, cartoonist (figure 2.5c)	coun-selling	collect family health history information	adult
(Zhou et al., 2018)	female	upper body, 2D, cartoonist (figure 2.5d)	coun-selling	alcohol use disorder counselling for adult	adult
(Bickmore et al., 2018)	female	upper body, 3D, cartoonist (figure 2.5b	health coach	mobile coaches for chronic condition management for adult	adult
(Bickmore et al., 2013a)	female	portrait, 2D, cartoonist	coun-selling	counselling about physical activity and food	adult
(Trinh et al., 2018)	female (3)	portrait, 2D, cartoonist (figure 2.5e)	consul-tant	physical activity and preconception care advisor, counselling Atrial Fibrillation patient	adult
(Bickmore et al., 2010a)	female	portrait, 2D, cartoonist	advisor	provide health information to inadequate health literacy patient	
(Niculescu et al., 2009)	male, female, androgy-nous	head shot, cartoonist, illustration (figure 2.3)	health guide	provide information of medical queries	adult
(Bickmore et al., 2010b)	female	headshot, 2D, cartoonist	consul-tant	promote antipsychtotic medical adherence for patient with schizophrenia	adult
(Utami et al., 2017)	female	headshot, 3D, cartoonist	empa-thetic partner	empathetic personal company for elderly people	adult

Table 2.3: Table shows details of the agent's appearance and different roles in the health domain.

(a) (Ring et al., 2016)

(b) (Schulman and Bickmore, 2009)

(c) (Wang et al., 2015)

Figure 1: Virtual Agent Counselor for AUD

(d) (Zhou et al., 2018)
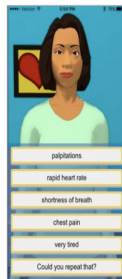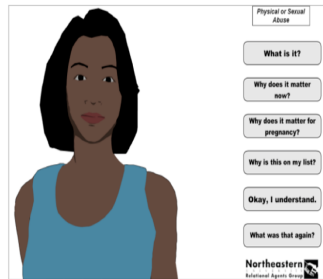
a. Carmen   b. Tanya   c. Gabby

(e) (Trinh et al., 2018)

Figure 2.5: Agent characters used in previous research in health domain.

Author	Gender	Embodiment	Role	Task	Age
(Khan and De Angeli, 2009)	female	headshot, cartoonist, 2D (figure 2.6a)	chatting partner	open ended chatting	adult
(Koda and Takeda, 2018)	female, male	upper body, 2D, japaneese anime (figure 2.6b)	chatting partner	formal conversation	teenager
(Silvervarg et al., 2012)	male, female, androgynous	headshot, cartoonist, illustration	chatting	chatting with the students	young, teenage
(Pelachaud, 2015)	female	none	conversation	can be a speaker or a listener in interaction	adult
(Castillo et al., 2018)	male, female	headshot, 2D, cartoonist	conversation	freely talk to the users	adult
(Bickmore et al., 2013b)	no gender	full body, 2D, robotics	museum guide	deliver information and talking with user	none
(Kopp et al., 2005)	male	upper body, 2D, cartoonist	museum guide	conduct small talk with visitors	adult
(Louwerse et al., 2005)	male, female	headshot, static, animated, cartoonist (figure 2.6d)	presenter	lecture video of Roman Empire	adult
(Cassell, 2001)	female	upper body, cartoonist	real estate	display and discuss about housing	adult
(Bickmore and Cassell, 2005)	female	full body, graphical, cartoonist	real estate salesperson,	answering user questions about properties	adult
(Payne et al., 2013)	male(4), female(4)	potrait, cartoonist, static	self service check out	none	adult
(Brahnam and De Angeli, 2012)	male (4), female(4), robot	headshot, cartoonist, static (figure 2.6c)	social presence	inquiry	adult, child(2)

Table 2.4: Table shows details of the agents' appearance and different roles in the entertainment domain.
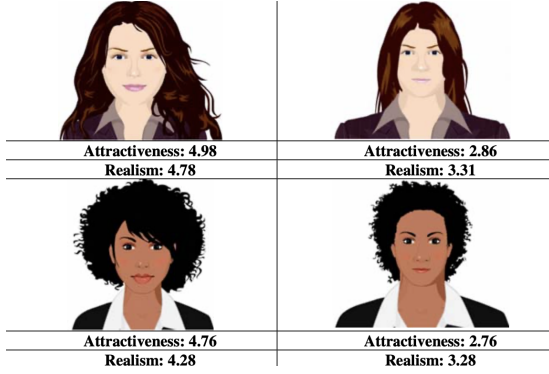
(a) (Khan and De Angeli, 2009)

(b) (Koda and Takeda, 2018)

(c) (Brahnam and De Angeli, 2012)
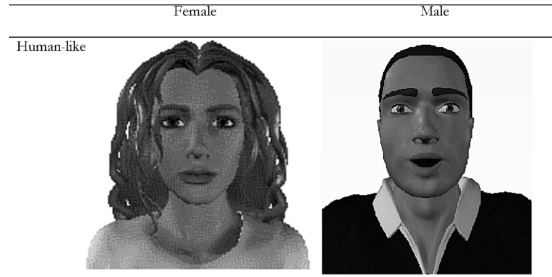
(d) (Louwerse et al., 2005)

Figure 2.6: Agent characters used in previous research in professional and entertainment domain.

In all these domains, numerous studies demonstrate that physical appearance influences the interpretation of actions and intentions, and thus appearance plays a significant role in the outcomes of social activities (Zebrowitz, 2018). Furthermore, research shows that the impact of physical appearance varies according to context (Zebrowitz, 2018). Forliz's (Forlizzi et al., 2007) explored the relationship between the visual features of embodied agents and the tasks they perform, and the social attributions in his research. The result of his research demonstrated a clear link between the task and form of the agent, and, also revealed that people often prefer that agents which possess the functions that are associated with gender stereotypes.

2.3 Gender Stereotypes in Virtual Embodied Agents

Gender stereotypes are acquired at a young age and can be very influential in human cognition and social behaviour (Martin and Ruble, 2010). According to the Social Role theory, these stereotypes tend to reflect on the underlying traits assigned to different genders, which are a direct result of observing different genders in their social roles (Eagly and Wood, 2012). These traits are mostly categorized as "communion", "agency" and "competence" traits where "competence" is also considered by some as a distinct trait (Bakan, 1966; Sczesny et al., 2018).

Communion trait includes qualities such as compassionate, warm and expressive; and is considered to be a stereotypical trait of the female gender. Agency includes qualities such as ambitious, assertive and competitive, which is regarded as the dominant male stereotypical trait. Competence (e.g., intelligent, innovative) has been conceptualized differently based on the domain and suggested to be correlated with both agency, warmth or even morality (Abele et al., 2016). When an individual has friendliness, trustworthiness from communion and assertiveness, competitiveness from agency, that is more inclined to competence. But agency and communion are negatively correlated to each other, Bakan suggested in his research. Agentic qualities like ambition, dominance, independence, or competitiveness have often been regarded as incompatible with communal purposes (Carlson, 1971). It is clear that "agency-oriented individuals experience fulfillment through their individual accomplishments and their sense of independence and separateness, whereas communion-oriented individuals experience fulfillment through their relationships with others and their sense of belonging" (Fiske, 2018). However, these stereotypes seem to be dynamic and have been changing based on the recent changes in our society.

One important indicator of dynamic stereotypes is the changing gender roles in our society since the mid-20th century, where we see progress towards gender-balance in previously male-dominated occupations (Statistics., 2019). A recent meta-analysis showed that, indeed, these changes in the society had an effect on stereotypical traits. It is found that competence equality in stereotypes increased over time, where male advantage over agency

seems to be diminished, and female advantage over communion is unaffected (Eagly et al., 2020).

Following up on this recent finding of Eagly and colleagues, the research of this thesis will investigate the changing effects of stereotypical traits for gender in virtual agents.

2.3.1 Virtual Agents & Gender Stereotypes

Current work on virtual agents mirrors the gender stereotypes, where female-looking agents are mostly used in roles that are associated with communion traits. Female agents have been dominantly used in applications such as personal companions (Khan and De Angeli, 2009; Pelachaud, 2015), assistants (Tsiourti et al., 2014), health counsellors (Bickmore et al., 2018; Trinh et al., 2018; Bickmore et al., 2013a; Schulman et al., 2011), and mental health counsellors (Ring et al., 2016). Male agents seem to be dominantly used in applications that require agency or competence traits such as mentors (Baylor et al., 2003a; Baylor and Ebbers, 2003), educational applications (Kopp et al., 2005; Massaro, 2004; Wik and Hjalmarsson, 2009) or motivators (Baylor and Ebbers, 2003). A recent UNESCO report (West et al., 2019) reveals the same trend in commercial voice-based conversational agents, which was criticized for reinforcing the gender bias in social context.

Most studies or commercial products tend not to provide any explicit reasoning for gender-biased design choices (West et al., 2019). Moreover, few of these agents actually go through a design process that compares the relative advantage of using a particular gender, let alone systematically evaluating the assumptions for gender preferences. The studies for gender preference for older adults show a preference towards female characters (Tsiourti et al., 2014). However, the appearance of these characters seems to include more differences in the variables such as clothing, the colour of skin/eyes/hair, or realism to be considered as conclusive. Several studies showed that users expect different levels of behaviour and perceive agents differently based on their gendered appearances, but the results can often be conflicting. For example, researchers showed that young students preferred to listen to engineering advice from male tutors more than females even though both male and female agents delivered the same persuasive message about engineering (André and Pelachaud, 2010). Whereas, when it comes to presenting lectures to adults (Louwerse et al., 2005) and

medical lectures to young adults (Gulz et al., 2007), participants preferred female agents than males. These results suggest the context-dependency of these effects, as well as the influence of other factors than gender stereotypes.

Gender stereotypes seem to play an important role in initiating first-impressions and expectations from the agent based on context (Rosenberg-Kima et al., 2008). Yet, several other causes were suggested as contributing factors that paint a more complex picture. Some researchers suggest attraction for other sex plays a role, as users seem to prefer the opposite-sex in formal-conversation (Koda and Takeda, 2018) or motivational tasks (Baylor et al., 2003b). However, these preferences seem to be context-dependent. Other studies showed users tend to be influenced more from the same gender (Kim et al., 2007; Guadagno et al., 2011a). More detailed evaluations conducted by researchers on other stereotypical factors such as race, attractiveness, coolness or age along with gender where they concluded that visual appearance and presence play important yet complex interactions with the context and with each other(Baylor and Kim, 2004; Baylor, 2009; Khan and De Angeli, 2009; Louwerse et al., 2005).

Moreover, very few studies touch on the flexibility of gender identities and the concept of androgyny. One study on pedagogical agents showed that androgynous characters could be a novel way to have both genders represented and overcome the abusive behaviour female agents often encounter (Silvervarg et al., 2012; Silvervarg, 2016). Another study used multiple male, female and one androgynous character, namely Ruth, in the health application setting (Niculescu et al., 2009), where the perception of the androgynous agent seems to differ based on the gender of the user; when female indicated androgynous character more feminine, male indicated more masculine but when female indicated androgynous character more masculine, male indicated more feminine. Another approach is to use Male characters with feminine traits, or female characters with masculine traits. However, the agents were not perceived differently and did not reflect gender stereotypes (Gulz et al., 2007); which suggests the need for a more controlled design process to achieve androgynous characters. Moreover, all of these studies used cartoon-like characters which allowed changing the physical characteristics of the agent in a more controlled manner but sacrificed realism.

These social effects can be more substantial with the manipulation of cues presented by the agent, such as visual appearance, gestures, the tone of voice or linguistic style, all of which highly influences the users (Blascovich et al., 2002). In this study, we focus on the visual cues, specifically the facial appearance of the agents and its effect on triggering gender stereotypes. The next section will include the related work in this domain.

2.3.2 Visual Cues for Gender

Ruttkay (Ruttkay et al., 2004) categorized the physical appearance of a synthetic character into static and dynamic visual characteristics. Static visual characteristics include basic physical properties such as body-type, face properties, skin, hair and hair-cut, and clothes. Dynamic visual characteristics include characteristics that can change in time, such as facial expressions, body and hand gestures, postures and movements. Both of these static and dynamic characteristics seem to have a role in gender recognition.

The visual markers that triggers gender stereotypes in agent interactions can include: clothing (Haake and Gulz, 2008; Gulz et al., 2007), body types (Jackson, 1992), hair style (Ueki et al., 2004), use of cosmetics (Gulz et al., 2007) as well as dynamic behaviors such as gaze (Kulms et al., 2011), emotional expressions (Halberstadt and Saitta, 1987) and gestures (LaFrance, 1981). Minimal visual exposure of most of these attributes can also effect the perception of age (Burt and Perrett, 1995; Enlow and Hans, 1996), attractiveness (O'Toole et al., 1998), personality traits (Sutherland et al., 2015) and even trustworthiness (Oosterhof and Todorov, 2008).

In terms of facial appearance, studies shown that eyebrows, face outline, nose (Brown and Perrett, 1993; Yamaguchi et al., 2013) and substitution of brows, chin, brows and eyes, and jaws (in ascending order) significantly alters the perception of gender when seen with the prototypical Caucasian male or female faces (Brown and Perrett, 1993). Eyebrows had more effect on studies with prototypical Japanese faces (Yamaguchi et al., 2013). Moreover, in isolation, it is found that eyes carry more information about gender than other parts of the face (Brown and Perrett, 1993; Roberts and Bruce, 1988).

However, the number of facial features to be considered, as well as the number of types for each feature and their interactions with each other, creates a complex space to evalu-

ate every possible composition. The standard approach for this issue is to systematically modify each facial feature and ask human participants for the assessment of the trait of interest. Even though this may not be an optimal solution for understanding the specific correlations, for the purposes of this thesis work, I will focus on the modification of static visual characteristics for the face.

2.3.3 Summary

From the literature on agent research, it is clear that a large number of virtual agent applications mirror the gender stereotypes. A great deal of research has been done on revealing and proving the stereotypes phenomena, but most of them employ agent designs that visualize 2D images, cartoonist models, gaming characters and binary gender characters for the agent research and application. However, the involvement of androgynous character and a more fluid notion of adopting gendered agent character to this research has yet to be unrevealed. The trend for changing gender stereotypes, changing gender roles and a dynamic understanding of the notion of gender has not been studied yet in virtual agent research yet.

Three significant issues standout as possible limitations from previous studies are -

1. A lack of a clear standard of what a virtual agent is, many studies use wildly different types of characters from stick figure to cartoonist, to illustrative/stylized to realistic and on another domain from 2D drawings and illustrations to 3d modelled and animated - all which add additional confounding factors that might bias the main gender results. Some studies do not even attempt to create rigorous input characters but instead, use 'found' characters from commercial systems with no regard for these confounding issues.

2. Many of these studies assume a binary notion of gender, which is either female or male.

3. Lastly and based on point 1, many of these studies given their female and male characters design, do not rigorously assure in the character design that the only difference between the two male and female characters is gender - that is from a clean study

perspective of measuring the only gender - they do not attempt to reduce all other factors (race, tallness, weight, hair colour) to assure to the best possible that users are seeing the same general agent other than gender difference.

A major contribution of this theses is to attempt to minimize these three limitations to better understand gender in Embodied Character Agents. Unlike the previous studies, this thesis aimed to create high fidelity human-like agent's character designs with a minimal amount of changes in the visual attributes, to eliminate the effect of various other factors that are influenced by the appearance of the agents by comprehensive research and analysis as well as it attempt to design with gender studies in mind. While opinions differ, it seems logical for virtual agent gender studies best to start with characters that are realistic humanoid designs (before moving to more ambiguous stylized designs). 3D realistic human-like characters would seem to allow for users to make their reception of the virtual agents closer to human gender reception. Realistic human-designs also, with like rigorous design guidelines that I have tried to deploy, would be easier to build where confound factors other than gender can be removed. This approach attempts to deal with limitations 1 and 3.

Different than previous work, this thesis introduced the concept of androgyny along with male and female genders to evaluate the stereotypical traits and roles and focused on individual agents with static facial cues for gender in human-like 3D agents. This deals with limitation 2. However, the aim of this thesis was not to evaluate the effect of each attribute on gender, instead to design good representatives for each category with minimal differences. This thesis is intended as an effort towards a progressive and inclusive approach for gender representations in virtual agents.

Chapter 3

Study 1: Gender Preferences for ECAs

In this chapter, I present a study that examines the effect of the gender-related physical appearance of embodied conversational agents (ECAs) during the face to face conversation with a human. The following sections exhibit the complete study process and its findings. The interaction scenario represents the variation of the agent's gender appearance involving a conversation with the human person. This study is the primary step to assess the ongoing gender stereotype perception of humans towards the agent. I should note that this first study is somewhat self contained, as it comes from published (poster) and presented work at the NeurIPS 2019 conference, in the WiML (Women in Machine Learning) workshop.

3.1   ECAs System Description

One issue with the published studies from this early research area (see Section 2.2.3) is that many studies use varying types of non standardized embodied agents from static images of characters to simple stick figures, to limited 2d animation, to cartoon limited 3d animation. These radically different interpretations of ECAs can affect and bias ECA gender studies with confounding variables. To counter this issue, I have used fully 3D advanced characters with game level realism, including facial and body expression and advanced rendering AI dialogue system for the study stimuli. It makes sense for these still early days of ECA gender research work to start with a high-end humanoid depiction with full human expression that more closely represent the humans before moving to other styled depictions. For that level

of depiction and dynamic movement the state of the art is using 3D animated characters. Therefore, I have used a socially-situated 3D virtual character human-like embodied conversational agent for this study, which can perform a set of verbal and non-verbal behaviour that allows for a real conversation with the interaction partner. These behaviours include facial expressions, gaze, head and body gestures, as well as verbal behaviours.

I use a modular system which is part of the multi year, multi PhD research work of my lab (ivizlab.sfu.ca) in cognitive based AI ECA research, in which I have contributed to with others. The work is documented in the following papers (Nag and Yalçın, 2020; Nag and Yalçın, 2019; Yalçin and DiPaola, 2018; Saberi et al., 2014; Karimaghalou et al., 2014; Arya et al., 2009; Arya and DiPaola, 2007; DiPaola, 1991; Turner et al., 2016; DiPaola and Yalçın, 2019; Yalçın and DiPaola, 2019; Nixon et al., 2018; Bernardet et al., 2017; Bernardet and DiPaola, 2015, 2014; DiPaola, 2013, 2009). For this study 1, I am using a variant of the system which uses as its character output engine, our lab's work tied to the USC (University of Southern California) Smartbody (Thiebaux et al., 2008) system. In Study 2 and 3, I create new characters that can be incorporated to our new Unity variant with modifications I have made.

Figure 3.3 displays the male and female embodied agent of the system that I used in study I.

Figure 3.1: Male ECA   Figure 3.2: Female ECA

Figure 3.3: Male and Female ECAs of the system.

To summarize the functionality of the system (Yalçin and DiPaola, 2018), the agent's behaviour can change and adapt to the user according to the state of the dialogue. While

the interaction partner is speaking, the agent enters the listening state to perform the text to speech. This listening mode will be activated via the speech and video input from the agent described in the perceptual module section. In this state, the agent is expected to provide proper backchanneling behaviour and personal feedback. After the speech of the interaction partner is completed, the agent will enter the thinking state. In this state, the agent will be finished gathering information from the perceptual model and start processing the speech input for generating a response. This response generation process will make use of the context of the dialogue and the emotional content of the message. Lastly, the agent will enter the speaking state, where the text to speech conversation based on the best AI-derived response is executed by lip synchronization with a high-quality voice. Therefore the audience sees that the ECA has simple, empathetic behaviour about and towards the user via face-to-face conversation.

3.2   Gender Preferences of ECAs

I have adopted a quantitative method to collect data from participants who have evaluated the gender-related appearance of the male and female embodied conversational agents in the academic counselling context. The ECAs 3D realistic appearance (see Section 3.1) with full facial, eye and body movements, real-time lip synchronization speech with a high-quality dynamic voice and other behaviour expressions made them natural and realistic to the subjects. Following the background on recent research that indicates gender stereotypes (see Section 2.2.3) in various roles and domains, I examine these effects for embodied conversational agents in the educational field. For gender stereotypical traits, I followed the well-studied personality and expertise variables (Gulz et al., 2007) to evaluate the user's preference for the agent.

The hypotheses of Study 1 are as follows:

S1_H1: Participants would be more likely to have affirmative preference during the academic counselling session if the appearance of the ECAs indicates as female.

S1_H2: Participants would be more likely to have adverse preference during the academic counselling session if the appearance of the ECAs indicates as male.

This study employs a quantitative research method to evaluate the agents' preference based on stereotypical gender traits to test the hypotheses.

3.2.1 Experimental Design

This study used a within-subjects design as the same participants tested all conditions. The first independent variable (IV) is the gender of ECAs, which has two levels: Male and Female ECAs. The second IV is the gender of participants, which has two tiers: male and female. The dependent variables (DV) are the personality and expertise of the agents. I have counterbalanced the conditions to reduce the chances of the order of conditions adversely influencing the results. All participants undergo condition M and Condition F. However, the order in which they receive differs; subjects were split into two groups randomly having the equal number of issues where one group was tested with condition M, followed by condition F, and the other was tested with condition F followed by condition M. The data analysis was performed by two way 2 (ECAs gender: male or female) *2 (participant's gender: male, female) mixed ANOVA with repeated measures. In a within-subject design, carryover effects can confound an independent variable's impact and affect outcomes and research results. They can pose a threat to validity.

Participants

Through snowball sampling and personal contacts, I have recruited participants from the graduate and undergraduate students of Interactive Arts & Technology at Simon Fraser University in Vancouver, Canada. It was incredibly challenging to recruit the participants because I needed equal number of male and female participants. People under eighteen were not included in the survey. The total number of participants is 80, where almost half (38) are female, and a half (38) are male, though 2 people declared their gender as other. Participant's age was between 18 to 40; the average age was 23.

Figure 3.4: Video clip recording setup.

Materials Design

Each participant watched two audio-visual clips that contain a back and forth conversation between ECA and the human person (figure 3.4). The clips are made out of the previous section (figure 3.1). The camera was only on the ECA, but the human user can be heard talking with it (figure 3.5). Their conversation topic was about graduate student life at SFU, such as how the user is dealing with academic study and regular life if any trouble or depression is going on. The ECAs were trying to counsel the user by providing empathetic suggestions, advice, and information if needed.

I have prepared two video clips for the study stimuli (condition M and condition F), where each stimulus condition contained one video clip.

- Condition M: The male embodied conversational agent and a human person (student) talk.

- Condition F: The female embodied conversational agent and a human person (student) talk.

Each video clip length was around two minutes. The video clips' aspect ratio was 4:3 at a size of 480 pixels wide by 360 pixels with high-quality audio. Figure 3.4 represents the instrumental setup environment of recording the video clips. The interaction scenario

Figure 3.5: Final video clip screen setup.

includes a face-to-face conversation between the agent and a human interaction partner, similar to a video conference where an audio speaker was used to record the clear voice quality and 360 video camera to film the video. I used the same setup for making the second video of talking with a female agent. However, I did not show the person's presence in the final video screen that is used in the study so that the participants can only focus on the ECA's activity (figure 3.5). Otherwise, it might distract them from concentrating on one point in the middle of ECA and a person's action on video. Participants can only see the portrait of the agents in the video clips. Except for their voice and gender appearance, all other aspects: conversational materials, gestures and expressions, everything was similar for both agents. The female agent has a female voice and a male voice for the male agent.

Gender preference traits were evaluated using the variables; helpful, natural response, comfortable (DV) used in various studies (Gulz et al., 2007). Each trait was calculated according to the total scores of associated adjectives on a 5-point Likert scale. The questionnaire was presented in random order for each participant. The total number of participants in each stimulus condition was 72 (34 male, 34 female and 2 others gender); they were assigned to two stimulus conditions by counterbalancing.

Procedure

The experiment was conducted by an electronic survey prepared in the Survey Monkey (www.surveymonkey.com) platform (see appendix B). Participants that qualified the inclusion criteria were automatically included in the study. Informed consent was taken from the participants that qualified for the study and were informed about the procedure before accepting the survey. Then, participants went through the survey process in the following order:

1. First page explained the study purpose, procedures and confidentiality issues. Once participants had consented to participate, the site tested whether their browser could display the video clips and prompted the participants to adjust their volume to a comfortable level.
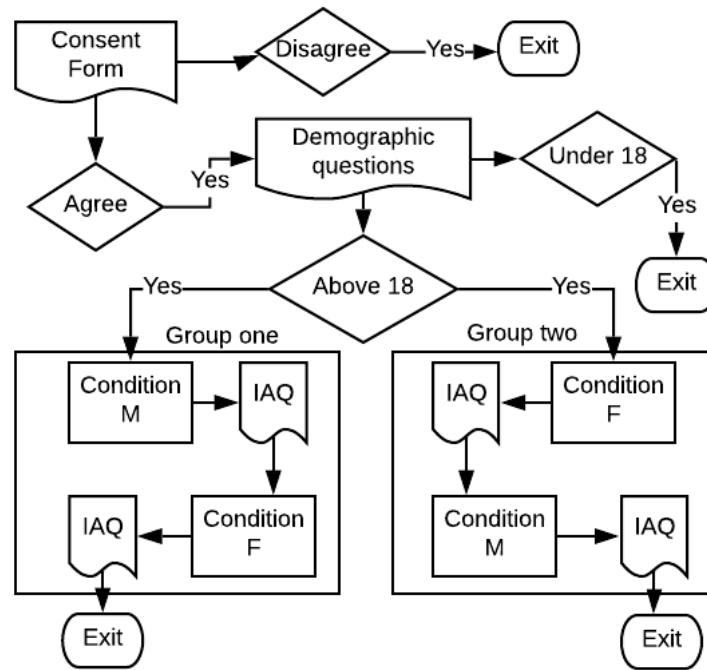

Figure 3.6: User study procedure.

2. The next screen represented a short description of the research to make the participants a bit familiar with the study and the demographic questions, which include

43

participants' age and gender. However, if a participant selected a period under eighteen, it would redirect the disqualification page and exit from the survey. When they selected any age range above eighteen, it would take them to the main survey page.

3. Participants were then presented with the first video clip in the sequence of stimuli. They were asked to watch the video at full screen with audio or headphones to get a better feeling of voice interaction and visual interaction.

4. After watching this video, the questionnaire page appeared. Participants selected their five-point Likert scale choice option of the five questions followed optionally by writing down their opinion in a comment box.

5. The survey interaction was designed in a way that the second clip appeared logically after completing the required IAQ responses. Later the questionnaire associated with the second video appeared after watching the clip. Participants performed a similar task for both conditions.

6. Lastly, notifying a 100% completion rate redirected to exit from the survey website.

However, based on informed consent, the participant could quit the survey at any stage. Thus, all participants in each group were exposed to the same clips' sequence regardless of their specific responses. All 72 participants undergo condition M and condition F. The first video clip displayed condition M and the second one displayed condition F for 36 participants during reverse order for the rest 36 participants. Hence, group one contained half of the total participants and group two contained the rest half.

3.2.2 Results

From 80 initial participants, only 70 completed the study and included in the evaluation. Among 10 participants who did not complete the study, 2 participants quit after completing 29%, 3 participants quit after completing 57%, 3 participants quit after completing 86%, and 2 participants denied consenting. I excluded those participants who mentioned their gender as "other" category because there were only 2 people in that category. Therefore, the final analysis was conducted upon 70 data containing half (35) male and half (35) female

44

participants responses. On the other hand, 44 male participants replied to the open-ended questions, but 38 female participants responded to the open-ended questions.

Eventually, I have processed the quantitative data of the study in SPSS from a two-way mixed ANOVA with repeated measures. Sphericity cannot be tested because there are only two levels (male and female) in the repeated measure. As a result of Mauchly's test in SPSS, the value is 1, the chi-square is 0, and the significance value is blank. Additionally, the Post Hoc test is not required as there are only two groups.

Personality Trait

The agents' personality is derived from the agents' empathetic gesture, such as helpful gesture, natural response, and comfortable attitude towards the user.

Comfortable score: The main effect of gender on comfortable scores was significant, $F(1,68) = 3.991$, p=.050, np^2=.055. This indicated that the comfortable level of conversation at which the participant's choice was measured is ignored; male and female participants' rating choice is similar. There was a significant main effect of comfortable conversation, $F(1,68)$=14.495, p<.001, np^2=.176. This indicates whether the rating came from male or female participants; the two comfortable types' ratings are significantly different. The interaction between comfortable and gender was non-significant, $F(1,68)$=2.406, p=.126, np^2=.034. This effect tells that the comfortable level of both male and female ECA's conversation has a similar effect on male and female participants. The pairwise comparisons for the main effect on engaging are corrected using Bonferroni adjustment. The test exhibits a significant difference p <.001 between feeling comfortable of male agent conversation and female agent.

These findings indicate that the agent's gender is influenced to see the comfort of conversation differently in men and women. From figure 3.7, the graph clearly shows that women seem to rate female agents' conversation more highly than men and others. However, all participants rated female conversation as more comfortable than males regardless of gender.

Helpfulness score: The difference across two ECA's helpful attitude was no significant $F(1,68)$= 2.994, p=.088, np^2=.042 and the helpfulness scores between groups was non-significant, $F(1,68) = 1.874$, p=.176, np^2=.027, in participants gender. However, the inter-
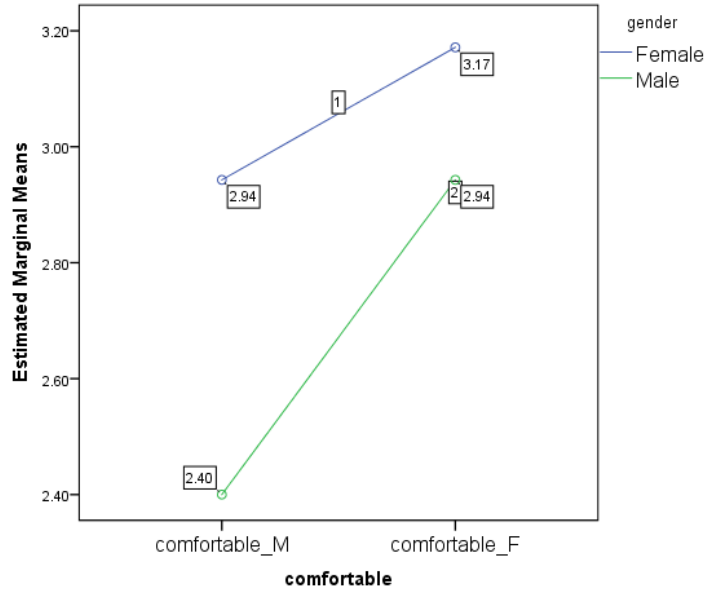
Figure 3.7: Male and female participants preference on feeling comfortable of male female agent conversation, comfortable_M contains the scores of participants rating for condition M and comfortable_F contains the scores of participants ratings for condition F

action between helpful attitude and gender was non-significant, $F(1,68) = .025$, $p = .875$, $np^2 = .000$. Following up, this interaction indicated that the response type of both male and
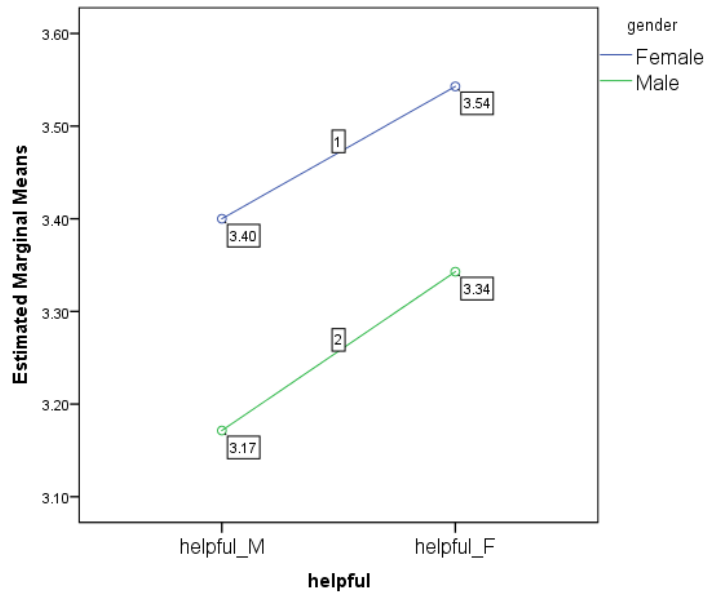

Figure 3.8: Male and female participants preference on helpful attitude of male female agent conversation, helpful_M contains the scores of participants rating for condition M and helpful_F contains the scores of participants ratings for condition F

female ECA's conversation had a similar effect on participants' gender. Bonferroni pairwise comparison test correction indicated a significant difference p<.001 between the helpful attitude of the male agent and the helpful attitude of a female agent. Figure 3.8 shows that participants found the female agent more helpful than a male agent.

Natural response score: There was a significant difference across two ECA's response type, $F(1,68)= 10.288$, p=.002, $np^2=.131$ but the natural response scores between groups was non-significant, $F(1,68) = .484$, p=.489, $np^2=.007$, in participants gender. However, the interaction between natural response and gender was non-significant, $F(1,68) = .548$, p =.462, $np^2=.008$. Following up this interaction indicated that response type of both male and female ECA's conversation had a similar effect on participants gender. Bonferroni pairwise comparison test correction indicated significant difference p<.001 between the natural response type of male agent and response type of female agent. Figure 3.9 indicates that participants felt female agent's response was highly natural than male agent.
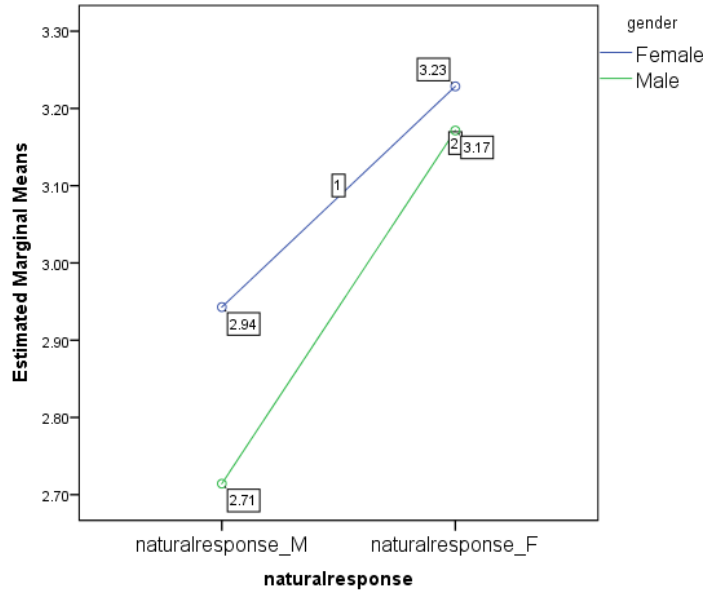

Figure 3.9: Male and female participants preference on natural response level of male female agent conversation, naturalresponse_M contains the scores of participants rating for condition M and naturalresponse_F contains the scores of participants ratings for condition F

Expertise Trait

The knowledge and experience of the defined area and ability to make an engaging conversation determine the agents' expertise on their role.

Interested score: There was a significant difference across two ECA's interested type, $F(1,68) = 22.015$, p<.001, $np^2=.245$ but the interested scores between groups was non-significant, $F(1,68) = 3.694$, p=.059, $np^2=.052$, in participants gender. However, the interaction between interested and gender was non-significant, $F(1,68) = 1.172$, p =.283, $np^2=.017$. Following up, this interaction indicated that the response type of both male and female ECA's conversation had a similar effect on participants' gender. Bonferroni pairwise comparison test correction indicated a significant difference p<.001 between the interest of getting a male agent and interest in getting a female agent. Figure 3.10 indicates that participants were more interested in getting a female AI agent assistant than a male.
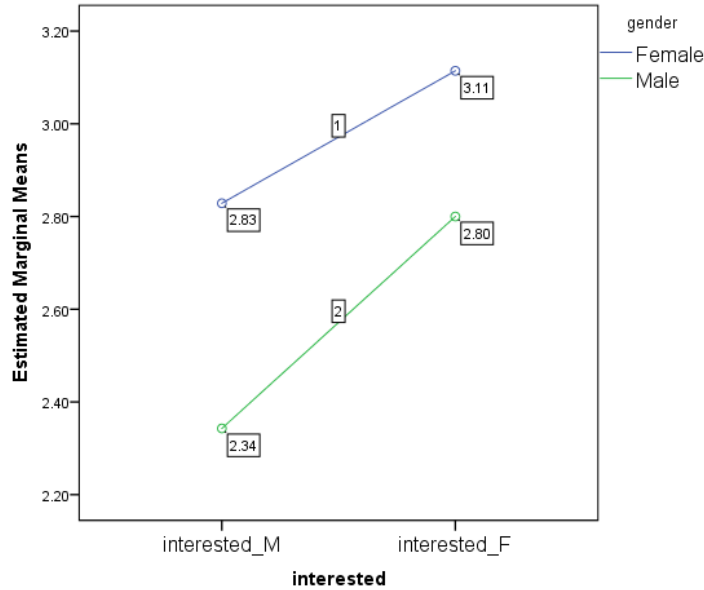

Figure 3.10: Male and female participants preference of interest level of male female agent conversation, interested_M contains the scores of participants rating for condition M and interested_F contains the scores of participants ratings for condition F

Engagement Score: The main effect of gender on the engagement scores was non-significant, $F(1,68 =8.814$, p=.004, partial eta square (np^2) =.115. This indicated that the engagement factor at which the participant's choice was measured is ignored; the rating

choice of male and female participants is similar. There was a significant main effect of engaging, F(1,68)=16.48, p<.001, np^2 =.195. This indicates whether the rating came from male or female participants; the two engagement types were significantly different. The interaction between engaging and gender was non-significant, F(1,68)=.878, p=.352, np^2=.013. This effect tells that the engagement level of both male and female ECAs conversation has a similar effect on male and female participants. The pairwise comparisons for the main effect on engaging are corrected using Bonferroni adjustment. The test exhibits a significant difference of p<.001 between male agents' engagement and female agents' engagement.

These findings indicate that the agent's gender influenced the conversation engagement differently in men and women. From figure 3.11, the graph clearly shows that women seem to rate female agents' conversation more highly than men and others. However, all participants rated female conversation as more engaging than males regardless of gender.

3.2.3   Discussion

Unlike some previous study which used recorded voice and a static figure illustration of an agent, the agent used in this study applied text to speech and speech to text to reply to the
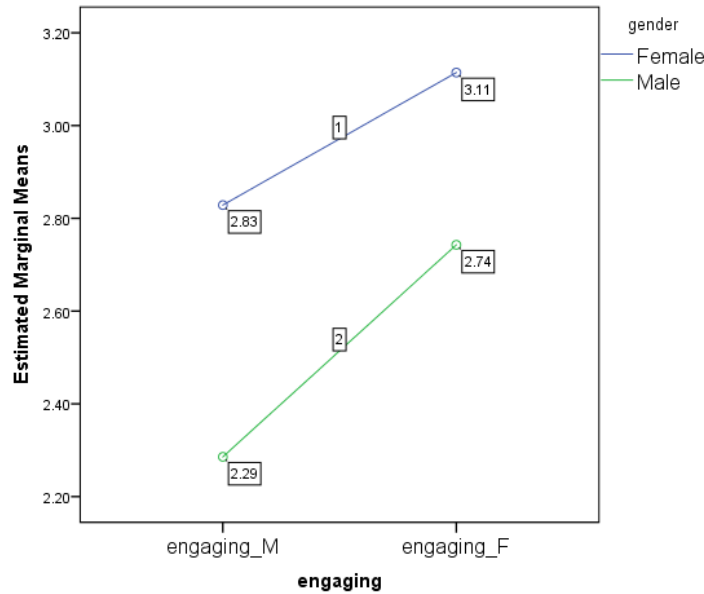

Figure 3.11: Male and female participants preference on engagement of male female agent conversation, engaging_M contains the scores of participants rating for condition M and engaging_F contains the scores of participants ratings for condition F.

49

user with its voice through lip synchronization to depict natural human talking. Both the female and male voices were of the same quality and execution; they came from Cereproc, which is known for their industry-leading high-quality dynamic text to speech voices. The high-quality real-time lip synchronization for each word and matching empathetic facial and body gestures made their appearance more natural compared to studies that use simple non-realistic characters or non-character voice-only agents. However, some aspects of the stimuli limited the possible conclusions of the study.

The results indicated a significant difference in how people responded to a male and female embodied conversational agent when presented with their interaction behaviours. The hypothesis S1_H1 predicted that participants would be more likely to have a positive interaction when ECA gives a female voice. The hypothesis S1_H2 expected participants to be more likely to have a negative attitude when the ECA presents a male appearance with voice. The results supported these two hypotheses.

Male vs Female Agent Preference

Participants felt the female response was more natural and comfortable than the male, even though the facial expressions, gaze, head and body gestures and dialogue were the same for both agents. On the other hand, from the responses, male and female participants' decision patterns across the helpful and engagement ranking are almost similar. Regardless of gender, the agents seem useful to the participants. Participants mentioned they are more willing to get a female agent as a personal advisor than a male agent. These findings showed a similar pattern as Baylor's study (Baylor, 2009). Even though participants mentioned both agents are helpful, they would prefer to have a female personal advisor.

From the above research findings, it can be said that the physical appearance (Massaro, 2004) and voice (Atkinson et al., 2005) was a factor in favour of a particular gender of agent or assistant, which indicate some of the current trends. Most voices on the digital assistant and other technologies are female. Scientific studies and psychological research influenced the decision to make these voices female. A professor from Stanford University, Clifford Nass, explains: "It's much easier to find a female voice that everyone likes than a male voice that everyone likes. It's a well-established phenomenon that the human brain is

developed to like female voices." Current digital assistant and previous research have proven that people like female chatbots or voice bots even if they are non-embodied, but this study revealed that the phenomenon is also as significant for realistic human-like embodied AI bots with synchronized multi-modal voice, face and body gesturing (including with empathy modelling).

The outcome can be the consequence of the visual cues (visual appearance) or social lines (voice) of the agents or by both. The agents have both physical appearance and vote in this study. Therefore, it leads to a critical analogy that has influenced participants' preference towards a specific gender, or both factors have the influence.

Design Implications

The study's scenario was just a two-minute video, not a real-time discussion with participants; therefore, participants might have felt different about more extended or other scenarios. If participants could have a real-time conversation with both ECA, they might develop a better or distinct sense of the agent's empathetic behaviour rather than watching video clips.

One potential concern of the study is the gender limit of the agent and participants. The study includes only male and female characters, no gender-neutral or androgynous characters, limiting the stereotype-related outcome. If an androgynous agent was also presented along with the male and female agents, participants' choice might differ or have a different opinion. Besides, participants' gender might affect the possible outcomes, as 99% of participants were from binary gender, only .01% of non-binary people took the study.

The visual appearance of the character emerges another potential drawback in terms of graphic design. Though all the verbal and non-verbal behaviours were the same for both agents, there were few differences in the visual appearance between the male and female characters. By analyzing participant's comments (see appendix C) to the open-ended questions, I have figured out these differences that can play a role into influencing participants' perception towards the agents. One of the significant differences between the agents is their skin colour; the male agent has a warm beige skin tone, whereas the female agent's skin colour is matte brown or tan. Additionally, their hair colour is also different;

male agents' hair colour is dark brown, whereas a female agent has blond hair colour. However, not a significant factor but their clothes colour is not the same as well. These design deviations need to overcome to create a balance between the agents' visual design.

And the five-point Likert scale might limit people to choose their options. The preference that participants showed for female ECA compared to males should probably be attributed to the scenario.

Lessons learned applied to study 2 and 3

From the above discussion, it is clear that the study results warrant further exploration, which could focus on potential differences in gender perception. By repeating this study in real-time where participants can talk to the ECA face to face like video calling and comparing the results, it would be possible to determine whether real-time discussion lessened the effect of the stimulus conditions, for example, by encouraging participants to reason from the facts and ethical principles. Additionally, real-time conversations between participants and ECAs can confirm the reliability of participants' perception of realism across different gender agents.

While this is an important direction and our advanced ECA system is capable of this level of AI interaction, one of the drawbacks of the study 1 research is that it does not have a proper balance of the visual design between male and female agents. Though all the verbal and non-verbal behaviours were the same for both agents, there were some differences in the visual appearance between the male and female characters. It was felt that this limitation might affect fair results on gender alone, and in noting that, looking at previous work, it appeared to be a limitation of many studies, which became a strong motivation to create rigour and balanced gendered characters for study 2 and 3 where all other confounding variables other than gender appearance are limited or removed.

Additionally if the agent's appearance indicates non-binary/gender neutral and more participation from non-binary people can be taken care of, that will help to get a conclusive output from any gender. Besides, additional information on open-ended responses from participants to shed light on their exposure to the embodied agent in various contexts has allowed us to understand better the factors influencing participants' decisions. Taking these

into account could improve a future study. As mentioned in design implications (see Section 3.2.3), if an androgynous agent was also presented along with the male and female agents, participants' choice might differ or have a different opinion. Therefore, I have added work by introducing an androgynous agent character in studies 2 and 3. It should be noted that, there can be other possibilities to create non-binary/gender-neutral character as well as agent such as robots, animals and different approaches. Since my research is on human-like agents including the expressions and gestures, I have limited the agents to human-like characters for this thesis. Including males, females, androgynous agent characters in the studies will strengthen the study evaluation and a possible conclusion.

Future work could isolate other attributes like social and visual cues that may distract from how the character is perceived, mostly along gender lines. These could include using an androgynous embodied AI character (differing voices) and male and female ECA and an equal number of transgender participants along with binary gender. However, Study 1 showed that it is essential to study visual and social cues separately first to understand their influence on user experience individually prior to turning any overall conclusion. So I deemed it important to start with visual cues first and gather clean data in this complex emergent area for this thesis work (Study 2 and 3) before moving on to interactivity and dynamic social cues.

With the results and lessons learned from Study 1, the next chapters will incorporate these lessons in Study 2 and 3, including validating rigorous high-end 3d characters created for gender reception and introducing the new and complex area of androgynes.

Chapter 4

Study 2: Character Design and Evaluation

In this chapter, I first present the methodology and hypotheses for Study II and III, as they are related, and best discussed together. For the rest of this chapter, I present the details of study II and then in the following chapter (Chapter 5), I present the details of study III including a major discussion of it results.

The main research of this thesis (called Study II and III) consists of a two-stage methodology to test the thesis's hypothesis: firstly, an iterative design process to create agents that have gender representative (male, female, androgynous) facial features. Then evaluating these agents based on the stereotypical gender traits and roles. The methodology is presented in the following two sections. The first stage, Study II, includes the design of the male, female and androgynous characters. The second and final stage, Study III, presents the evaluation study with the final characters from Study II, to test the main hypothesis of this thesis.

Following the background on recent research for changing gender stereotypes, changing gender roles and a dynamic understanding of the notion of gender (see Section 2.3 to review), this studies will examine these effects for virtual agents that show static facial cues for gender. The hypotheses for studies II and III are:

S23_H1: Following the recent findings of Eagly and colleagues (Eagly et al., 2020), the communion traits would be rated significantly higher for the agents that are perceived as female, followed by androgynous and male agents.

S23_H2: Following the recent findings of Eagly and colleagues (Eagly et al., 2020), the agency traits would be rated significantly lower for the agents that are perceived as female, followed by androgynous and male agents.

S23_H3: Following the recent findings of Eagly and colleagues (Eagly et al., 2020), the competence traits would not be significantly different between agents that are perceived as more female, androgynous and male.

Plenty of researches have shown that communion, agency and competence are the fundamental components of social perception (e.g., (Abele and Wojciszke, 2007; Fiske, 2018; Diehl et al., 2004; Carrier et al., 2014)). Agency refers to an individual's thrive to be the superior of the surrounding environment, to assert the self and experience the pleasure of achievement and power, they prioritized self advancement. In contrast, communion refers to an individual's desire to maintain close relationship, social connection and get along with others (Carlson, 1971), they prioritized togetherness. Unlike agency and communion, competence refers to the capability of adapting with circumstances and environment by utilizing the observable and measurable knowledge, skills, abilities and personal attributes that contribute to improve thyself effectively (see Section 2.3 for more details about these stereotype traits).

4.1   Design Rationals

The initial stage of this research consists of the creation of 3D agents based on gender-related facial cues. We have designed human-like 3D characters to develop an embodied conversational agent, using a 3D character design tool, Adobe Fuse (Adobe, 2017).

The design process consists of an iterative evaluation method with design students that include qualitative and quantitative methods. For both iterations, we changed the static facial attributes that are found to contribute to gender perception in humans (see Section 2.3.2). These attributes include cheeks (males with more prominent/angular cheekbones), jaw and chin (males with long/square jawline and chin), lips (females with fuller lips), neck (males with Adam's apple), eyebrows (male eyebrows straighter), and hair (males with short hairstyle than males).

Unlike the previous studies, we aimed to achieve character designs for female, male and androgynous characters with a minimal amount of changes in the visual attributes to eliminate the effect of a variety of other factors that are influenced by the appearance of the agents (see Section 2.3.1 for a review). We removed the effects of clothing, skin colour, hair colour, eye colour, height, body type, and different facial expressions. We also created agents with facial symmetry, taking into account its role in sexual attraction (Shamekhi et al., 2018).

4.2   Agent's character design process

I have followed an iterative design process to design embodied agent's characters.

4.2.1   First Iteration of Character Design

Based on the same basic face, but differing by masculine and feminine attributes (Gulz et al., 2007). I have taken into account the concepts presented in (Shamekhi et al., 2018) that symmetry, averageness and sexual dimorphism are the three main factors to make a face attractive. The more symmetrical a face is, then the more attractive it is perceived to be. In order to exhibit the human-like naturalism in the faces, some detailed elements are added in all characters such as visible freckles around the nose and cheeks, wrinkles around the outer corner of the eyes and other parts of the faces, deeper under eye area than the face surface and slight dark circles shades.

We started with an average representative character for the initial iteration and manipulated the aforementioned facial attributes to create initial examples for female, male, androgynous agents. The differences between the first round of characters are summarized in Table 4.1, and the characters can be seen in Figure 4.1. Two androgynous agents were created by mixing the facial properties of males and females. After creating the characters' initial set, we evaluated their gendered appearances to get first impressions from design students.

Properties	M1	M2	F1	F2	A1	A2
Eyebrows	straight	straighter M1	arched & thick	F1	M1	F1
Eyelashes	regular	M1	fuller lashes	F1	F1	M1
Cheek	high & angular	M1	round & low	F1	F1	F1
Jaw/Chin	square & long	F1	round & short	pointier F1	M1	F1
Lips	narrow	M1	fuller	F1	F1	M1
Lip color	no	no	red	red	red	no
Neck	wide & adam's apple	M1	narrow	F1	wide	F1
Hair	short	M1	mid-length	long	short	A1

Table 4.1: Table shows details for the first round of character design. Rows show the attributes that were changed according to the intended gender of the agent. Columns show the names of the different agents created. Columns with names that start with M shows agents expected to have Male-gendered attributes, F for Female and A for Androgynous. If agents have the same design attribute with other agents, their name is presented instead of describing the attribute.

Experimental Design

To evaluate the perception of gender of the designed characters, participants were exposed to assess some male, female, androgynous embodied agents through an electronic survey. Androgynous is someone who presents as not incredibly feminine or masculine. It's not gender identity, how they express themselves.

Participants: Participants were 15 graduate and undergraduate students (8 male, 6 female, 1 identified as 'other') working in a design school and part of an intelligent agent visualization lab. The average age of the participants was 25 (SD=1.06). All participants know 3D virtual agent and previously participated in a character design process or class.

Materials and Design: At the beginning of the survey, participants were asked two demographic questions, age and gender. We have also asked if they are familiar with 3D characters and have done any courses or work related to 3D modelling to assume their character design knowledge. Because if anyone has this kind of experience, they can identify the difference between attributes manipulation of different images; otherwise, it may not seem much different to general people. The mandatory question asked about their thought of
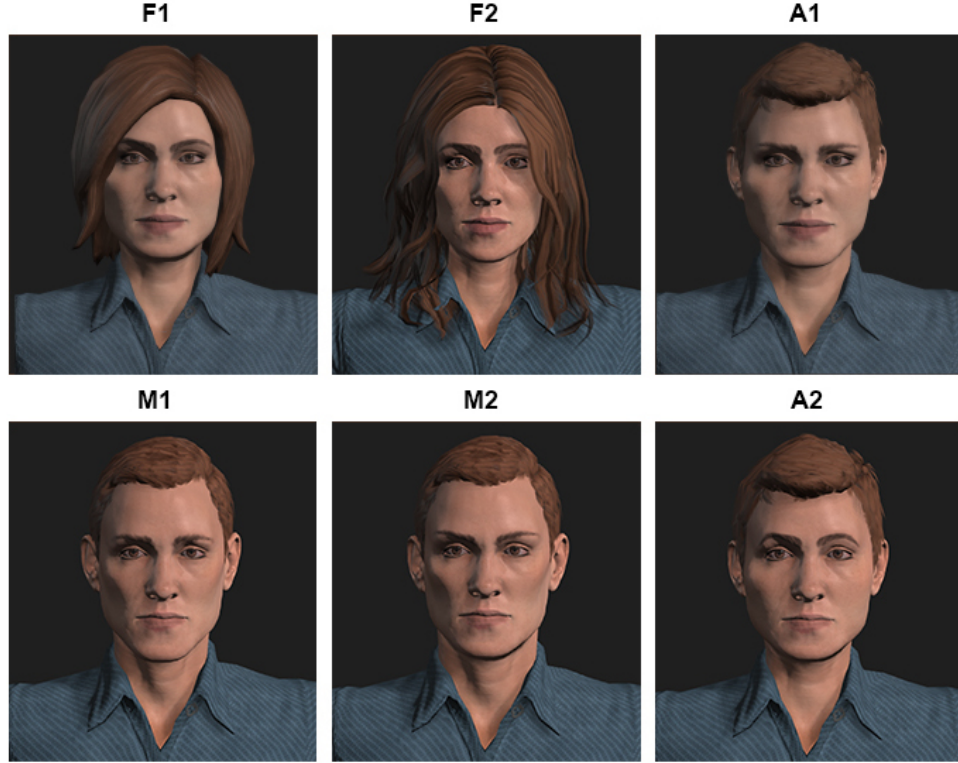
Figure 4.1: Variation of male, female and androgynous characters appearance (First iteration).

the gender of the characters from its visual appearance along with an open-ended comment box asking about their opinion about the character's design.

Procedure: The study was conducted in the Survey Monkey (www.surveymonkey.com) platform (see appendix D.1). Informed consent was taken from the participants before beginning the study. Participants were asked about the demographic information (age, gender) and their familiarity with 3D characters and modelling. After this stage, participants were informed about the study procedure and told about the meaning of the term "androgynous." Participants were then shown the images of created characters in random order and asked to rate each character based on their gender perception with a slider with 20 equally distributed points (0: male, 10: androgynous, 20: female). Each page also included a comment section that asked about their opinions on why they think agents should be rated. Participants finished the survey with an additional comment screen for further comments about the agents' overall design.

Results: The results showed that participants have generally rated the male characters as male (M= 2.1, SD= 3.04) and female characters (M= 18.4, SD= 3.07). In comparison, the rating for androgynous characters was averaged closer to males with more deviation (M= 7.9, SD= 5.60) for the two versions. Hence, I analyzed the qualitative data from participants' comments to improve our androgynous agents. One of the main attributes that the participants highlighted was hair. The test showed a tendency for perceiving the androgynous character as a male due to the short hair. The hairstyle also played a vital role in rating females. Participants often commented on the prominent check bones for considering the androgynous character as a male, but almost none commented on the jawline. The comments also showed no mention of the neck (Adam's apple) in the qualitative evaluations. The agents' eyelashes did not seem to provide enough difference to change the participants' idea about the agents' gender. Table 4.2 shows the detailed results for all agents-

	F1	M1	A1	A2	F2	M2	A3
Sum	279	31	118	118	273	30	83
Median	18.6	2.06	7.86	7.86	18.2	2.14	5.92
SD	1.6	3.15	5.8	4.75	2.90	3.00	6.03

Table 4.2: Result of the first iteration of character design process.

Discussion: The results show male and female characters' preferences were close to our male and female agents' expectations. However, the differences between the two male characters and the two female characters were not significant. Therefore, we picked the two agents with the least variation (F1 and M2) to include in the second design iteration. However, the ratings and comments for rating the androgynous characters as males made us explore further into redesigning this character. One participant commented, "the softer eyes and lips bring feminine qualities, while the stronger eyebrows, jawline, larger nose brings more masculine qualities and hairstyle gives me a more masculine impression" for the androgynous character. Moreover, as the presence of Adam's apple and eyelashes were overlooked, we standardized these attributes between agents and removed them from our design decisions for the next iteration.

4.2.2 Second Iteration of Character Design

To further refine the representative androgynous character's design, we iterated over the agents we picked from the first design cycle. First, we standardized some overlooked features by removing Adam's apple, wide neck, and fuller eyelashes for all agents. Next, considering hair might change some variables such as cheek shape by blocking the visual cues, we decided not to include long hair and instead use a swept-back style that would be ambiguous in terms of its length. Moreover, we created multiple androgynous characters by having several attributes from both female and male agents. Among the number of designs we created, we included the agents with the attributes represented in Table 4.3, which are shown in Figure 4.2. The same procedure and participants for the first iteration were used (see appendix D.2).

Properties	F1.2	F3	A1.1	A1.2	A1.3	A1.5	A1.6	A1.7	A1.9	M3
Eyebrows	arched & thick	F1.2	M3	M3	M3	M3	M3	F3	F3	straight
Cheek	round & low	F1.2	F3	F3	F3	M3	F3	F3	F3	angular
Jaw/Chin	round & short	F1.2	M3	F3	F3	F3	F3	F3	F3	square
Lips	fuller lips	F1.2	F3	F3	M3	F3	F3	F3	M3	narrow
Lip Colour	red	red	no	red	no	no	no	red	no	no
Hair	mid-length	swept back	F3	F3	A1	F3	F3	A1	F3	short

Table 4.3: Table shows details for the second round of character design. Rows show the attributes that were changed according to the intended gender of the agent. Columns show the names of the different agents created.

Figure 4.2: Variation of male, female and androgynous characters appearance (Second iteration).

Results: Results showed that participants preferred the agent A1.1 as a better representative of androgyny (M= 11, SD=3.5). This character's formation consists of jaw, chin and eyebrows like the M1 character, checks and lips like the F1 character. The representative male (M=2, SD=2.9) and female (M=17, SD=3.4) characters were still reported as male and female. Table 4.4 shows the detailed result of the second iteration for all designs.

	A1.6	A1.7	A1.9	M3	A1.1	A1.2	A1.3	A1.5	F2	F3
Sum	159	187	67	31	172	224	42	223	258	240
Median	10.6	12.46	4.46	2.06	11.46	14.93	2.8	14.86	17.2	16
SD	3.41	6.3	4.90	2.90	3.58	3.57	2.93	4.40	3.58	3.25

Table 4.4: Result of the second iteration of character design process.

Discussion: After the final evaluation, we picked the characters F1.2, M3, and A1.1 to be used in the stereotype evaluation study to represent examples of female, male and androgynous agents. The differences between female and androgynous characters only include eyebrows and jaw. Our aim was not to evaluate the effect of each attribute on gender, but rather to design good representatives for each category with minimal differences. We did not want to have perfect representatives with exact changes but good representatives with minimal modification. There can be many other different versions that we could come up with, but we only needed exemplars for this study.

Chapter 5

Study 3: Evaluation of Stereotype Traits

Study III was conducted to examine the effect of gender stereotypical traits and roles based on the agents' gendered appearance. The research hypothesis has been described at the beginning of chapter 4 (see Section 4 to review hypothesis).

For gender stereotypical traits, I have followed the well-studied communion, agency and competence variables (DV) (Eagly et al., 2020). The research of Beyer (Beyer, 2018) found that there are no changes over 16 years from 1996 to 2012 on the gender pay gap due to the low awareness of occupational segregation. Their study showed a list of male and female-dominant roles and neutral roles over this time period. For instance, nursing and counsellor as female dominating, editor/reporter and financial manager as neutral and physician as dominating male roles. Therefore, for stereotypical roles, we have analyzed the U.S. labour force statistics (Statistics., 2019; STATISTICS, 2020) for the current number of workers in different domain and roles that are typically female/male-dominated roles and mentioned in the previous research. We chose 2 female-dominated roles as: healthcare-support (men=184, women=1287), assistants (men=137, women=1932); 2 male-dominated roles: exercise coaches (men=122, women=35), financial advisors (men=296, women= 152). Moreover, we chose 2 representative roles to represent gender-balanced occupations such as arts, design, entertainment, sports, and media occupations sectors (men=965, women=806). As for artists (men=48, women=35) and news reporters (men=40, women= 37) roles. We

tried to assign a specific mode of entertainment category, so chosen storyteller for that and news reporter to represent gender-balanced roles.

5.1 Experimental Design

The first independent variables (IV) are participants' gender, which has two levels: male and female, and the gender-related representation of the agents, which has three levels: male, female and androgynous. The dependent (DV) variables are communion, agency and competence characteristics of the agents and agent roles described above. The study used a within-subject design. A 2 (gender of participants: male and female) x 3(gender of agents: male, female and androgynous) mixed design has been conducted to test our hypotheses.

5.1.1 Participants

A total of 78 participants were recruited from Amazon Mturk. Only the participants over 18 years old with English as a first language were allowed to participate in the study; no other recruitment restrictions were applied. Excluding the incomplete data, the final number of participants who participated in the study was 72 (41 male, 31 female, 0 other). 10, 43, 15, 2, 2 participants were from 18 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 65 age range respectively. So, the majority participants were from 25-34 age ranges.

5.1.2 Materials and Design

Three portrait images of 3D adult human-like characters (i.e., male, female, androgynous) were used for this study, which was finalized from the previous study (see section 4). Each agent is presented in randomized order, resulting in 6 different combinations in order. Participants were counterbalanced.

Gender stereotypical traits were evaluated using the communion, agency and competence (see Section 2.3 to review) variables (DV) used in various studies (Eagly et al., 2020). Each attribute was calculated according to the total scores of associated adjectives on a 5-point Likert scale. Communion trait includes: affectionate, compassionate, sensitive, and emotional. The Agency trait has: ambitious, aggressive, courageous, and decisive. Competence

trait includes: creative, intelligent, innovative and organized. Adjectives were presented in random order.

Stereotypical roles were evaluated with 2 female-dominated roles (health coach, personal assistant), 2 male-dominated roles (exercise coach, financial advisor) and 2 gender-balanced roles (storyteller, news reporter) that are evaluated by using 5-point Likert scales. Roles are presented in random order.

Each agent's gender perception was evaluated with a 5-point Likert scale slider (female=0, male=5, androgynous=3). This question included a definition of androgyny.

5.1.3 Procedure

The experiment was conducted on Amazon Mechanical Turk, linked to a survey prepared in the Survey Monkey platform (see appendix E). Participants that qualified the inclusion criteria were automatically included in the study. Each participant was given 1$ (USD) upon completion of the survey. Informed consent was taken from the participants that qualified for the study and were informed about the procedure before accepting the survey. Then, participants went through the survey process in the following order:

1. Participants were presented with a short description of the research and the demographic questions.

2. The portrait images of the agents were presented in random order with gender stereotypes and roles questionnaires.

3. Participants were asked to rate the gender of the agents. The same portrait images of the three agents were presented with the gender questionnaire in a randomized order. An optional section asked about their decision on the agents' gender.

5.2 Results

From 78 initial participants, only 72 completed the study and included in the evaluation. The result of Study III was processed via SPSS. The value of communion is formed by summing up affectionate, compassionate, sensitive, and emotional input. Competence values

are the summation of creative, intelligent, innovative, organized, and agency summation of ambitious, aggressive, courageous, decisive values.

5.2.1 Agents' Gender Scores:

I first evaluated the perception of the gender of female, male and androgynous designed agents. The results showed that the designed agents were indeed perceived significantly differently based on their gender and consistent with their desired appearance ($F(2,140)=$ 148, p<.001), where male and female participants scored agents' genders similarly (p>.5). There was a significant main effect of agents' appearance on identifying the agent's gender ($F(2,140)=$ 148, p<.001, np^2=0.680). But there was no significant main effect of gender on identifying the agent's gender overall (p>0.5); both participants performed similarly for male (M=13.10), female (M=13.49) and androgynous (M=13.39) characters. In addition, there was no significant interaction between agents' appearance and gender in terms of identifying agents' gender (p>0.5). The descriptive statistics showed that the result of males (M=4.46, SD=.69), female (M= 1.08, SD= 1.29) and androgynous (M=2.94, SD=1.49) are almost close to the actual scale. From Figure 5.1, we can see that both males (M= 2.82) and females (M= 2.84) participants performed quite similarly overall in identifying agents' gender.

5.2.2 Gender Stereotypical Traits

The main effect of communion (w=.956, x2(2)=2.723, p=.256), agency (w=.988, x2(2)=.695, p=.706) and competence (w=.986, x^2(2)=.825, p=.662) the assumption for sphericity was not violated. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality. Therefore, the F-value for communion's main effect (and its interaction with between-group variable, gender) does not need to be corrected for violation of sphericity.

Communion Trait: For the communion scores, there was no significant main effect of agents appearance, F(2,140)=.87, p=.42, np^2=.01); or gender, F(1,70)=3.48, p=.07, np^2=.05, on communion scores overall ; male (M=13.870) and female (M=12.452) participants performed quite differently overall. In addition, there was no significant main effect
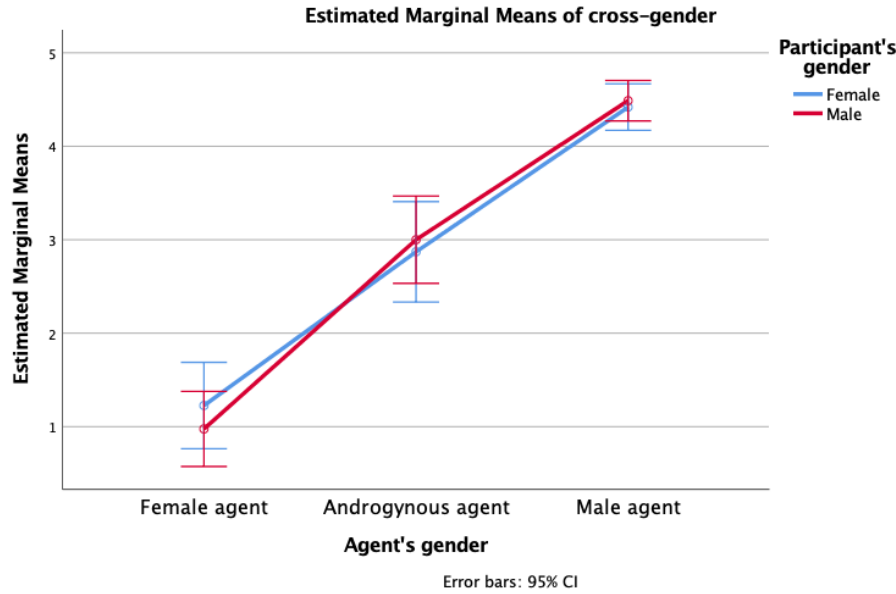
Figure 5.1: Agents gender related appearance and its correlation with the gender score based on participants gender.

of gender on communion scores overall, $F(1,70)=3.476$, $p=0.066$, $np^2=0.047$; participants performed quite differently for male (M=12.939), female (M=13.327) and androgynous (M=13.216) characters. However, there was a significant interaction between agents' gender and gender of the participants for communion scores, $F(2,140)=2.68$, $p=.03$, $np^2=.05$. Descriptive statistics showed that males (M=13.08, SD=3.64), female (M=13.44, SD=3.62) and androgynous (M=13.25, SD=3.48) characters got similar ratings overall from participants. Male participants rated female agent character (M=14.17) higher in communion than female participants' ratings (M=12.48). A similar trend was seen in male characters' ratings for male participants (M=13.98) than for female participants (M=12.97). Figure 5.2 shows the trends for each category.

In the graph, the blue line represents the female participants, and the red line represents the male participants. From the graph 5.2, we can see those female participants found an androgynous character, most communion and male character least communion. In contrast, male participants found female character most communion and androgynous character least communion.
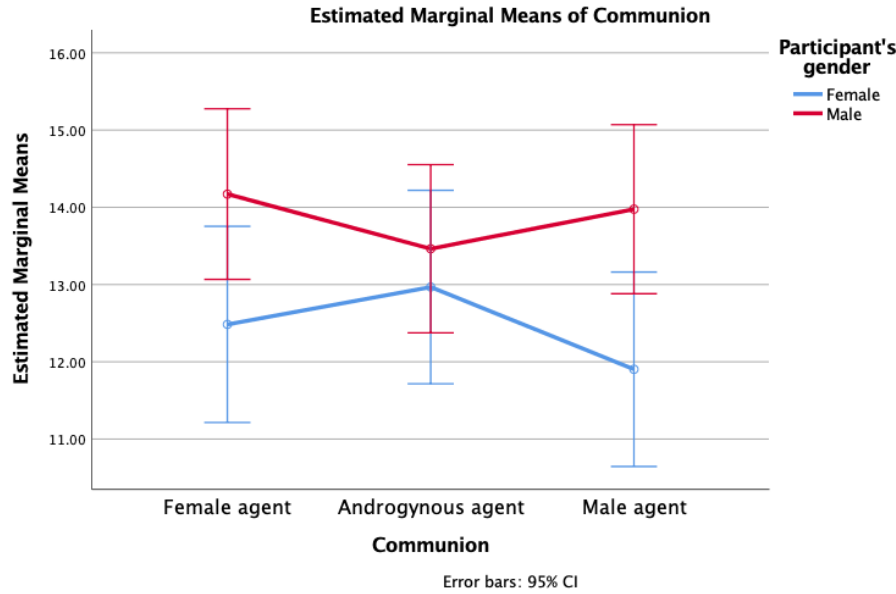
Figure 5.2: Agents appearance on communion scores.

Agency Trait: There was no significant main effect of agents'gender on agency scores, participants' gender on agency scores or interaction between agents' gender and participants' gender (all p>.05). Both females (M= 13.45) and male (M= 13.97) participants performed quite similarly overall, and they performed similarly for males (M= 13.63), female (M= 13.74) and androgynous (M= 13.77) characters. Descriptive statistics showed that female character got higher ratings (M= 13.82, SD= 2.77) than male (M= 13.64, SD= 2.72) and androgynous (M= 13.78, SD= 2.62) characters overall from participants. Male participants determined female agent character (M= 14.34) more agency than male (M= 13.71) and androgynous (M= 13.85) character. Female participants showed a different pattern; they rated all the agents quite similarly, male (M= 13.55), female (M= 13.13) and androgynous (M= 13.68).

Figure 5.3 shows that female participants found the androgynous character most agency and female character least agency. In contrast, male participants found the female character most agency and male character least agency.

Competence Trait: No significant main effect of agents' or participants' gender over the competence trait scores was found (p>.05), as well as the interaction between agents' gender and participants gender (p>.05). For the competence result, there was no significant

68
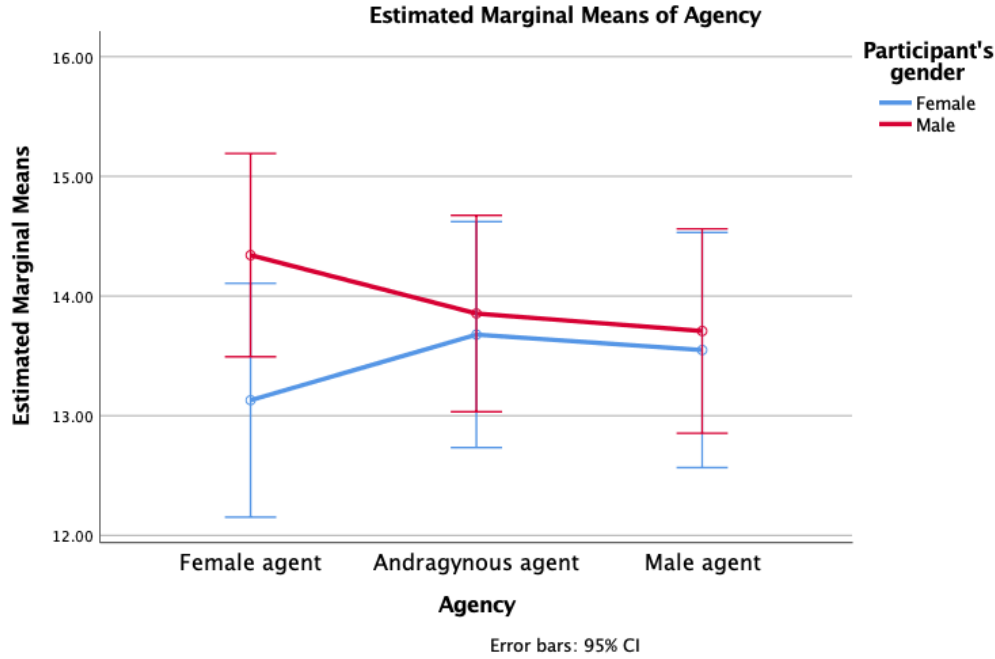
Figure 5.3: Agents appearance on agency scores.

main effect of agents appearance on competence scores, gender on competence scores and interaction between agents appearance and gender in terms of competence scores, p>.05; both male (M=14.748) and female (M=14.452) participants performed quite similarly over-all. Both participants performed similarly for male (M=14.579), female (M=14.588) and androgynous (M=14.632) characters. Descriptive statistics showed that male (M=14.59, SD=2.335), female (M=14.61, SD=2.860) and androgynous (M=14.652, SD=2.819) character got similar ratings from participants.

From figure 5.4, female participants found androgynous character most competence and male character least. In contrast, male participants found male characters' most competent and female character least.

5.2.3 Gender Roles

News Reporter Role: There was a significant main effect of agents' gender on news reporter role scores, $F(2,140)= 4.37$, p=.01, np^2=.06; both female (M=3.73) and male (M=3.63) participants performed quite similarly. In addition, there was no significant main effect of gender on news reporter scores overall, $F(1,70)=0.39$, p=0.53, np^2=0.01; both
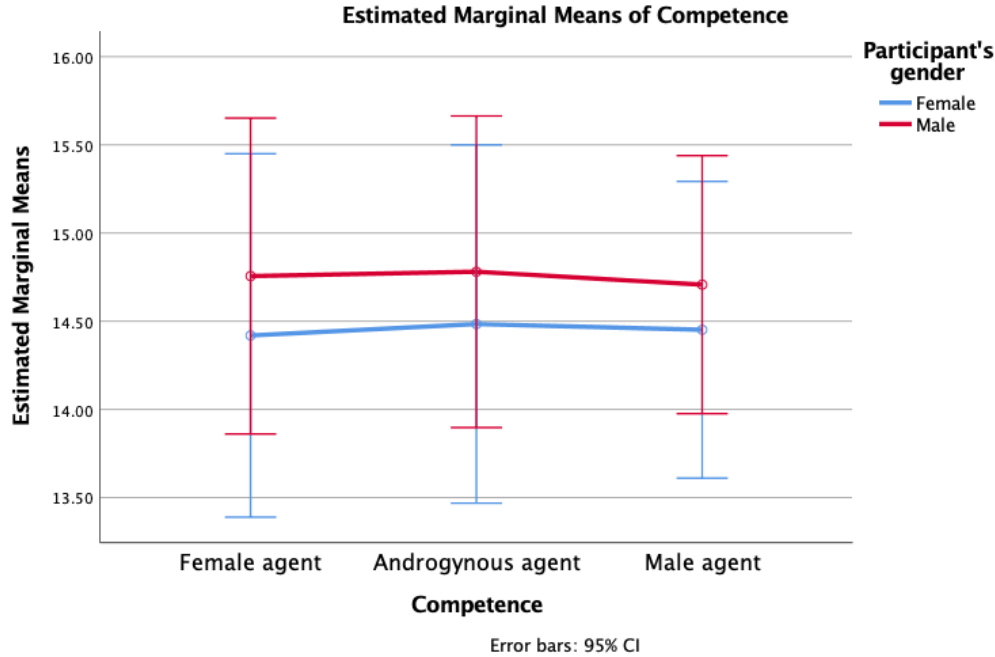
Figure 5.4: Agents appearance on competence scores.

participants performed similarly for male (M=3.70), female (M=3.86) and androgynous (M=3.49) characters. There was no significant interaction between agents' and participants' gender in terms of news reporter role scores, $F(2,140)= 2.018$, p= .14, $np^2= .03$. Descriptive statistics showed that female characters got higher ratings (M=4.86, SD=.79) than male (male=3.71, SD=.89) and androgynous (M=3.46, SD=1.01) characters overall from participants. Male participants determined female agent character (M=3.85) more like news reporter than female (M= 3.76) and androgynous (M=3.29) character. Female participants showed a similar pattern; they determined female (M=3.87) character more like exercise coach than male (M=3.65) and androgynous (M=3.68) character. Figure 5.5 shows the trends for each category.

The Anova result showed no significant main effect of agents' gender on a personal assistant, health counsellor, exercise coach, financial advisor, storyteller role, p>.05.

Personal Assistant Role: No significant main effect of agents' or participants' gender over the personal assistant role scores was found (p>.05), as well as the interaction between agents' gender and participants' gender (p>.05). For the personal assistant result, there was no significant main effect of agents appearance on individual assistant scores, gen-

70
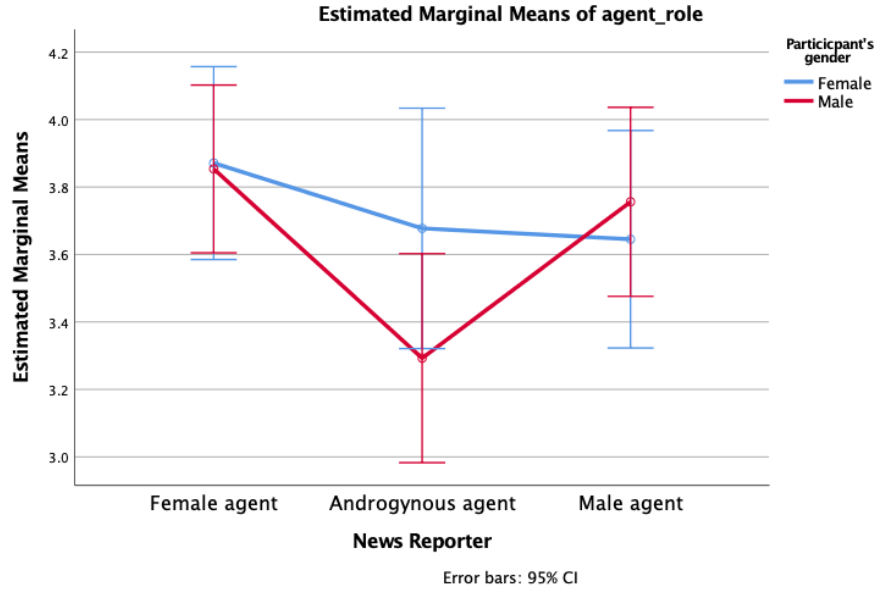
Figure 5.5: Agents appearance on news reporter role scores.

der on personal assistant scores and interaction between agents appearance and gender in terms of personal assistant scores, p>.05; both male (M=3.602) and female (M=3.545) participants performed quite similarly overall. Both participants performed similarly for male (M=3.539), female (M=3.583) and androgynous (M=3.599) characters. Descriptive statistics showed that males (M=3.54, SD=1.006), female (M=3.58, SD=0.884) and androgynous (M=3.58, SD=1.071) character got similar ratings from participants.

From figure 5.6, female participants preferred androgynous agents more as a personal assistant than other characters. In contrast, male participants showed the opposite preference.

Health Counsellor Role: No significant main effect of agents' or participants' gender over the health counsellor role scores was found (p>.05), as well as the interaction between agents' gender and participants' gender (p>.05). For the personal assistant result, there was no significant main effect of agents' appearance on health counsellor scores, gender on individual assistant scores and interaction between agents' appearance and gender in terms of health counsellor scores, p>.05; Male (M=3.333) and female (M=3.602) participants performed slightly differently. Both participants performed similarly for male (M=3.507) and female (M=3.531) characters but a somewhat different for androgynous (M=3.365)

71
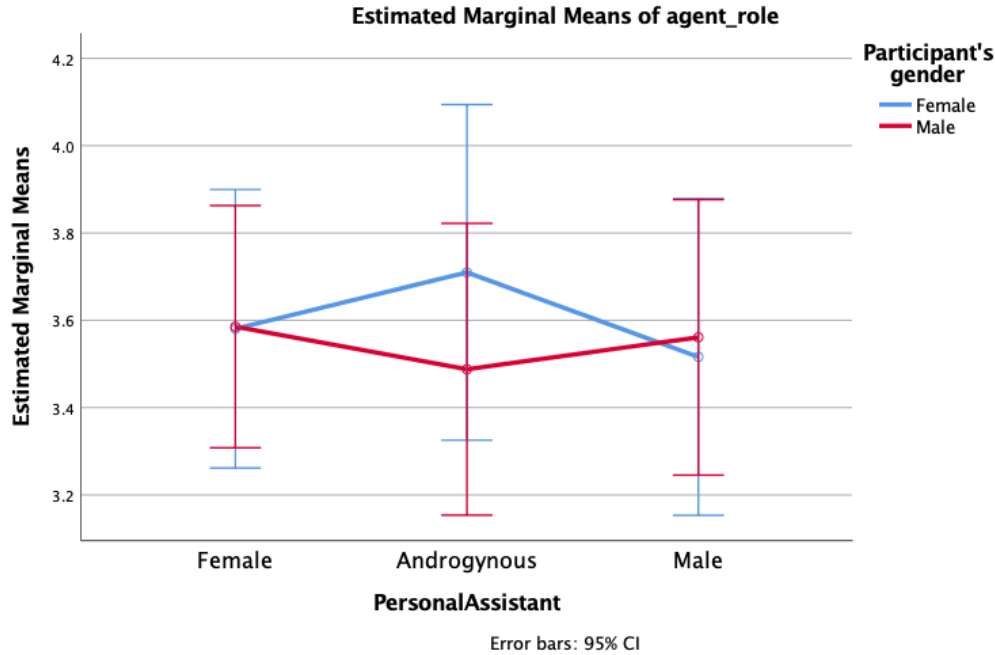
Figure 5.6: Agents appearance on personal assistant role scores.

character. Descriptive statistics showed that males (M=3.54, SD=1.006), female (M=3.58, SD=0.884) and androgynous (M=3.58, SD=1.071) characters got similar ratings from participants.

From figure 5.7, female participants preferred female character more for health counsellor and a male preferred male character. Both male and female participants showed less preference for an androgynous character than other characters.

Exercise Coach Role: No significant main effect of agents' or participants' gender over the exercise coach role scores was found (p>.05), as well as the interaction between agents' gender and participants' gender (p>.05). For the exercise coach result, there was no significant main effect of agents' appearance on health counsellor scores, gender on exercise coach scores and interaction between agents appearance and gender in terms of exercise coach scores, p>.05; Both male (M=3.355) and female (M=3.423) participants performed similarly. Both participants performed similarly for male (M=3.474), female (M=3.315) and androgynous (M=3.377) character. Descriptive statistics showed that males (M=3.49, SD=1.021), female (M=3.31, SD=0.959) and androgynous (M=3.39, SD=1.108) characters got similar ratings from participants.

72
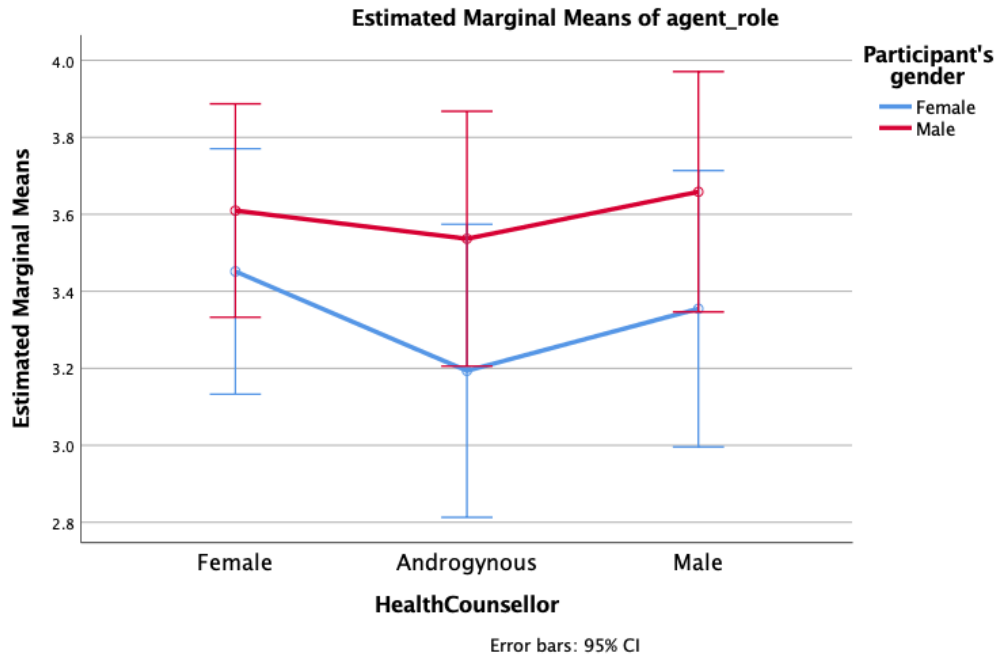
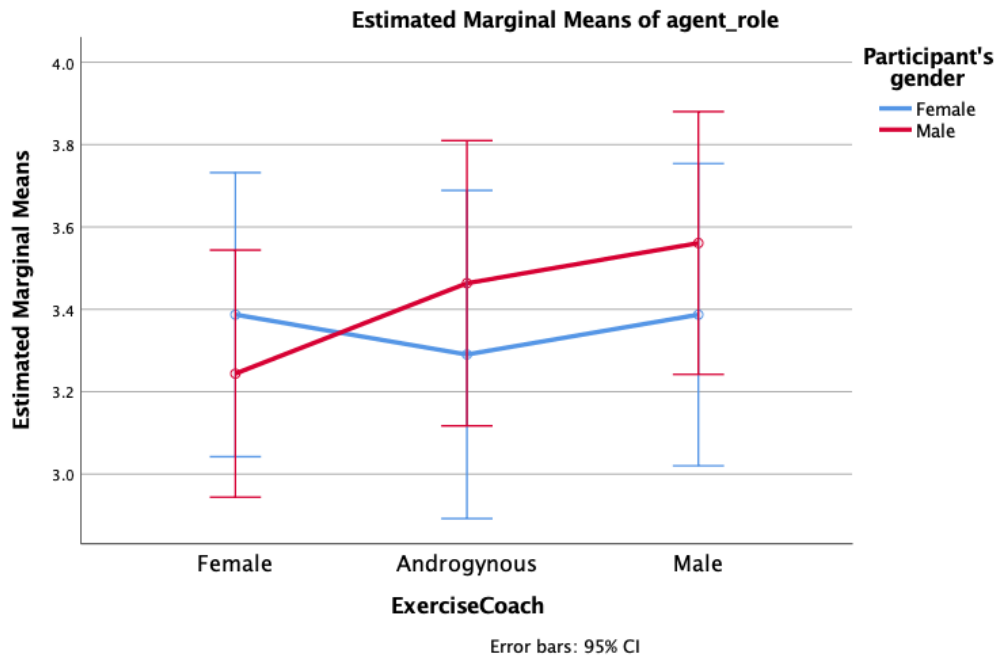Figure 5.7: Agents appearance on health counsellor role scores.

Figure 5.8: Agents appearance on exercise coach role scores.

From figure 5.7, female participants preferred male character more for exercise coaches than others. In comparison, male participants showed a similar preference for both male and female characters.

Financial Advisor Role: No significant main effect of agents' or participants' gender over the financial advisor role scores was found (p>.05), as well as the interaction between agents' gender and participants' gender (p>.05). For the financial advisor result, there was no significant main effect of agents' appearance on financial advisor scores, gender on financial advisor scores and interaction between agents' appearance and gender in financial advisor scores, p>.05; Both male (M=3.742) and female (M=3.805) participants performed similarly. Both participants performed similarly for male (M=3.871), female (M=3.785) and androgynous (M=3.664) character. Descriptive statistics showed that males (M=3.88, SD=0.821), female (M=3.78, SD=0.809) and androgynous (M=3.68, SD=0.962) characters got similar ratings from participants.
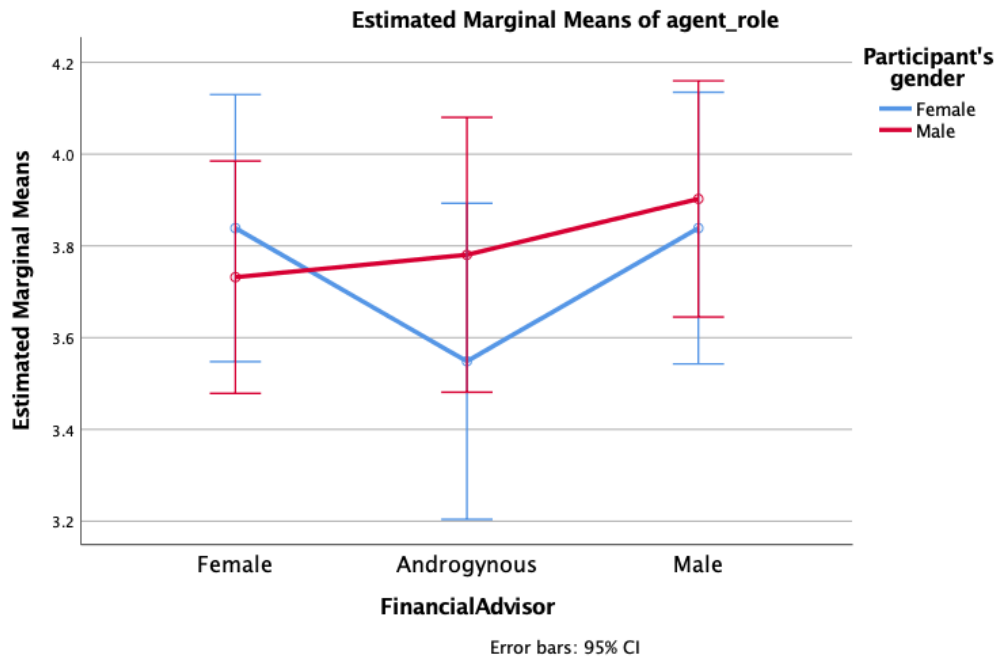
Figure 5.9: Agents appearance on financial advisor role scores.

From figure 5.9, female participants preferred androgynous characterless for the financial advisor than other characters. Whereas, male participants liked the male character more for financial advisor than other characters.

Story Teller Role: No significant main effect of agents' or participants' gender over the storyteller role scores was found (p>.05), as well as the interaction between agents' gender and participants' gender (p>.05). For the storyteller result, there was no significant main

effect of agents' appearance on storyteller scores, gender on storyteller scores and interaction between agents' appearance and gender in terms of storyteller scores, p>.05; Male (M=3.172) and female (M=3.415) participants performed quite differently. Participants performed differently for male (M=3.118), female (M=3.438) and androgynous (M=3.324) characters. Descriptive statistics showed that males (M=3.13, SD=0.020), female (M=3.47, SD=0.074) and androgynous (M=3.33, SD=0.113) characters got similar ratings from participants.
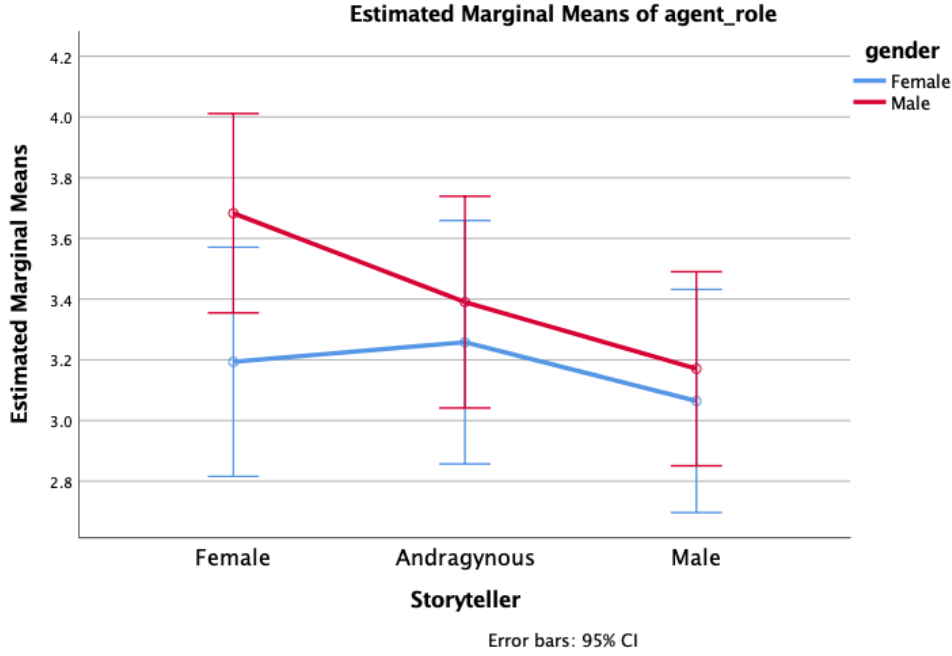
Figure 5.10: Agents appearance on story teller role scores.

From figure 5.10, both male and female participants preferred male characterless for a storyteller than other characters. Male participants liked female characters more for storyteller than other characters. At the same time, female participants preferred female characters more than male but less than androgynous characters.

5.3 Discussions

The results showed that this rigorous design process successfully generated representative male, female, androgynous agents with minimal differences. Unlike previous work, I evaluated the gender perception of our designed agents before evaluating the gender-based

stereotypes, used realistic-looking 3D characters instead of cartoon-like representations and included androgynous agents in the thesis studies. Results for androgynous agent showed a trend that positions the agents' scores between female and male agents. This is consistent with my hypotheses and previous studies (Silvervarg et al., 2012; Silvervarg, 2016) that suggested androgynous agents could be used to achieve a middle ground for the negative and positive effects of the male and female gender. As there are very few studies on androgynous agents, I hope that these design methodologies and results would encourage future work and further research in inclusive technologies.

The findings mostly point towards a gender-balanced view of stereotypical traits. I found that even though the participants successfully perceived the differences in the agents' genders, the communion, agency and competence traits mostly did not show differences between agents. This result further extends the meta-review findings by Eagly and colleagues (Eagly et al., 2020) on the rapidly closing gap in gender stereotypes. Similar to the thesis study results, Eagly and colleagues found no difference in the competence trait scores of males and females (S23_H3). Moreover, they found a trend towards gender equality in the agency traits with a diminishing male advantage and a female advantage on communion traits. However, competence and agency scores of this thesis showed greater gender equality. Male, female and androgynous agents were perceived similarly in agency and communion, rejecting S23_H1 and S23_H2.

This result could be due to the fact that in gender stereotype questionnaires for humans, the questions are asked about the general idea of 'femaleness' and 'maleness' rather than in relation to specific images for each gender. As I controlled for many variables of facial attributes in the design process, this might indicate that most of the stereotypical traits related to the agency could be more related to the collective bias in gender, rather than specific attributes of an individual. This may suggest that any gender representation could be used to embody different traits, and gender alone might not be a justified reason for choosing one agent over the other.

Moreover, a few of the participants also noted that "agency" implied a "living creature," "organism," or "alive," and therefore could not be considered as having agency. This per-

spective could also explain the reasoning behind having similar agency ratings regardless of gender. However, further examination is needed to conclude.

Another important finding in study 3 was the significant effect of the interaction between the participants' gender and the agents' gender in communion scores. Female participants' scores were significantly lower for all agents, where androgynous agents were scored above male and female agents. A counterpart of this effect was previously noted in other studies, where participants seem to favour their sex in communion scores (Eagly et al., 2020). We further see a reversed effect of the ratings on androgynous agent, where female participants' ratings for androgynous agent's communion scores are higher than average, contrasting male participants' lower-than-average rating for androgynous agents. This finding needs further investigation to be explained conclusively, as it was not previously studied in research on gender stereotypes.

The non-significant results of female and male-dominated roles also showed that the mostly non-explicit reasoning behind using female-gendered agents as "assistants," "health-care, "and male-gendered agents as "coaches" or "advisors" (West et al., 2019) is not supported. These results indicate that these roles could be used interchangeably, even with androgynous agents. I obtained a significant result only for the "news reporter" role, which was considered as a gender-balanced category. I saw that androgynous agents were rated lower for this category. I believe this finding would require a further inspection as there are no comparative results in human stereotype research on androgyny.

Chapter 6

General Conclusion

This final chapter summarizes the research contributions in this thesis. First, I reiterate the research objectives I presented in Chapter 1. I then describe my research contributions by outlining how I achieved each of my thesis goals from Chapter 1. Lastly, I discuss opportunities for future work for gender stereotypes in an embodied agent.

6.1 Thesis Problems

This thesis investigates the design and evaluation of embodied agents around gender stereotypes. Embodied virtual agents (ECAs) are computer-generated animated anthropomorphic interface agents which can engage a user in real-time, multi-modal dialogue, using speech, lip synchronization, gesture, facial expression, gaze, posture, intonation, and other verbal and nonverbal behaviours to exhibit the experience of face-to-face conversational interaction like human interaction (Cassell et al., 2000; Huang, 2017) (see Section 1.1 to review about ECAs). The overarching research problem of this thesis is: we do not know whether the gendered appearances of agents alone affects the perception of gender stereotypical traits or the preference over certain stereotypical roles. To more easily address this overarching research problem, I have broken my main problem into three sub-problems:

Research Problem 1: When only varying with binary gender-related appearance of embodied virtual agents (ECAs), how does the experience of the user change?

The findings in most virtual agent research and studies show extensive usage of female agents in commercial systems. This overuse of a specific gender in ECAs appears to give

a warning sign of the gender stereotypes issue (Khan and De Angeli, 2009; Pelachaud, 2015; Ring et al., 2016; Schulman and Bickmore, 2009; Wang et al., 2015; Silvervarg and Jönsson, 2011; Zhou et al., 2018; Bickmore et al., 2018, 2013a; Trinh et al., 2018; Bickmore et al., 2010b; Utami et al., 2017; Bickmore and Cassell, 2005). On top of that, l (Payne et al., 2013; Louwerse et al., 2005; Gulz et al., 2007; Koda and Takeda, 2018). While many previous studies show that users prefer female over male agents, in the study of Andre and Baylor (Baylor et al., 2003a; André and Pelachaud, 2010) audiences preferred a male tutor when it comes to technical or engineering subjects over a female tutor. However, the results from the more recent scientific literature on virtual agents show a changing trend on gender stereotypes in recent years (Eagly et al., 2020). Using realistic 3D ECAs, this thesis investigates these current gender stereotype trends by varying gender representation in user studies.

Research Problem 2: Regarding moving past binary distinctions of male and female, how does the more fluid notion of gender (or even gender neutral) affect a user's experience with ECAs?

The modern concepts of gender fluidity and androgyny have not been well studied in the virtual agent research community. The recent increase in awareness of non-binary gender categorization and movements makes it essential to approach androgynous agent character and male and female. Very little research has been done on providing guidelines to create an androgynous character for a virtual agent. Therefore, researchers need to investigate this area starting with how to represent the androgynous character's visual appearance for an embodied agent.

Research Problem 3: How do the gender-related appearance of male, female and androgynous agents impact the user perception in terms of gender stereotypes traits?

The dominance of using female cues in agents in assistant roles received much criticism (West et al., 2019) as these agents are thought to enforce social bias and do more harm than help in closing the gap on gender inequality (Specia, 2019). Moreover, recent research suggests that these gender roles are fluid, and the differences between gender stereotypes

are not as biased in our society. These stereotypes tend to reflect on the underlying traits assigned to different genders, which are a direct result of observing different genders in their social roles (Eagly and Wood, 2012). These traits are mostly categorized as "communion" and "agency," where "competence" is also considered by some as a distinct trait (Bakan, 1966; Sczesny et al., 2018). A comprehensive meta-analysis of Eagly et al. has shown a dramatic change in stereotypical traits and functions in the last several decades (Eagly et al., 2020). Therefore, it is important to reevaluate the user experience towards embodied agents in terms of gender stereotypes traits and roles (see Section 2.3 and 1.1 to review).

These above factors indicate the need for a revised look at gender cues' effect in virtual agents. The old assumptions and research need to be revisited in light of new evidence on the critical gap in gender stereotypes. Moreover, gender appearance on perceived stereotypical traits and gender roles has not been previously studied in virtual agent research.

6.2 Thesis Contributions

The contributions can be summed as follows. In an iterative design process, I created representative male, female and androgynous agents with as few differences as possible in their visual attributes. Second, I used these agents to evaluate the stereotypical assumptions of gendered traits and roles in virtual agents. The results showed that gender stereotypes are not as effective as previously assumed, and androgynous agents could represent a middle-ground between gendered stereotypes. I believe this finding would require a further inspection as there are no comparative results in human stereotype research on androgyny. As there are very few studies on androgynous agents, I hope that this design methodology and its results would encourage future work and further research in inclusive technologies.

Research Goal 1: Examine the effect of gender during conversations with an embodied conversational agent.

In chapter 3, I have examined if the user experience changes when presented with gender distinctions of embodied conversational agents (ECA). Though some previous research has shown gender stereotypes studying virtual agents, very few have used embodied agents that can perform face to face conversation in real-time through verbal and non-verbal be-

haviours to validate the trend. In my study, I have presented embodied conversational agents that are full of gestures, facial expressions and can continue an AI-enabled conversation with an interaction partner through voice back and forth. This study has shown that aspects of a virtual agent's gender appearance can influence moral or ethical consequences in gender-specific ways. Although subjects favoured females more than men overall, only men were significantly affected by the appearance factors, reacting positively to females in the F condition. The fact that males and females respond differently to changes in a character's visual presentation has design implications, impacting future systems created to facilitate interactions in counselling and other scenarios. If extraneous factors prove capable of subconscious influence, careful consideration of character presentation should be vital to a system's design.

Research Goal 2: Examine the user experience on perceiving the gender-related appearance of 3D characters with minimal differences in visual cues.

In chapter 4, I came up with the design and evaluation of male, female and androgynous characters. The design requirements for my problem emerged from study 1 (see Section 3.2.3 for a review) and related literature described in chapter 2 (see Section 2.2 for a review). One of the significant limitations of study 1 was not including an androgynous/gender-neutral agent to evaluate the gender stereotype assumption. Additionally, the difference between the agents' visual appearance indicated the added importance of a more balanced visual design, such as a better balance between the skin colour, hair colour, and clothing colour scheme. Most previous studies used cartoonist character, simple avatar, static image of 2D or 3D character, a stick figure of a character with a traditional chat-bot system to validate the gender stereotype assumptions. Additionally, most previous studies have adopted only binary gendered characters to validate the gender stereotype assumption. My main contribution is to create an androgynous character to diminish the gender stereotype boundary.

I came up with study 2 to diminish these design deviations by creating binary and non-binary 3D agents based on gender-related facial cues. I have designed human-like 3D characters to develop an embodied conversational agent, using a 3D character design tool,

Adobe Fuse (Adobe, 2017). The design process consists of an iterative evaluation method with design students that include qualitative and quantitative methods. For both iterations, I have changed the static facial attributes found to contribute to gender perception (see Section 2.3.2). Unlike the previous studies, my target is to achieve character designs for female, male and androgynous characters with a minimal amount of changes in the visual attributes, to eliminate the effect of a variety of other factors that are influenced by the appearance of the agents (see Section 2.3.1 for a review). I removed the effects of clothing, skin colour, hair colour, eye colour, height, body type, and facial expressions. I also created agents with facial symmetry, taking into account its role in sexual attraction (Shamekhi et al., 2018).

Research Goal 3: Examine the impact of the gender-related appearance of the agents on user perception in terms of gender stereotype traits and roles.

Chapter 5 presents a study to see the general population's point of view to correlate with the recent research for changing gender stereotypes, changing gender roles, and a dynamic understanding of the notion of gender. The three main categories of stereotypical traits (i.e., agency, communion and competence) and examples of roles were examined, supported by the recent statistics on the labour force. The study results supported the findings of Eagly and colleagues (Eagly et al., 2020) on the diminishing effect of gender stereotypes. This thesis study is intended as an effort towards a progressive and inclusive approach for gender representations in virtual agents. The non-significant results of female and male-dominated roles showed that the mostly non-explicit reasoning behind using female-gendered agents as "assistants," "health-care, "and male-gendered agents as "coaches" or "advisors" is not supported. The results indicate that these roles could be used interchangeably, even with androgynous agents. It should be noted that Eagly et al. (Eagly et al., 2020) conducted a meta-study; they did not run an actual analysis with a virtual agent to support their claim. I have employed their findings to the virtual agent research and conducted a comprehensive study using the high-end realistic agents I created in chapter 4 to evaluate their claim.

6.3 Limitation and Future Works

One of the drawbacks of the study 1 research does not have a proper balance of the visual design between male and female agents. Though all the verbal and non-verbal behaviours were the same for both agents, there were some differences in the visible appearance between the male and female characters. It was felt that this limitation might affect fair results on gender alone, and in noting that, looking at previous work, it appeared to be a limitation of many studies, which became a strong motivation to create rigorous characters for study 2 and 3.

In study 1 (and again in many previous studies), research was limited to binary gender, male and female agent. If an androgynous agent was also presented along with the male and female agents, participants' choice might differ or have a different opinion. The study's conversation scenario was limited to a two-minute video, not a real-time discussion with participants; therefore, participants might have felt different about a longer or different scenario. If participants could have a real-time conversation with both ECA, they might develop a better or distinct sense of the agent's empathetic behaviour rather than video clips.

In Study 2, only survey study was used to evaluate the characters' design. Interviewing study participants might open up some new findings rather than just looking into their comments from the open-ended responses. Also, the carry-over effect may encounter during the within-subject experiment of the study. In future work, I would like to incorporate more qualitative interviewing, especially for user reception to the androgynous agent, because this is quite a new area, and I do not have any clear theory or hypothesis given the lack of prior research. Additionally, I have designed an androgynous agent character to represent a gender-neutral human-like character, but there can be other approaches to create gender-neutral agents characters such as robots, animals and different approaches. Since the thesis is focused on the human-like agent, I have not explored other approaches to create a gender neutral character.

It should be noted that much of the later research was conducted in a lockdown based on the COVID virus, which limited our work to non-face to face approaches.

Study 2 and 3 used non-interactive techniques, even though we had a fully interactive AI ECA system. Our state-of-the-art current lab ECA system is capable of this, including with empathy modelling; however, it was deemed essential to start with non-interactivity first, control variables within this new line of research and with that data move on to face to face conversational interactivity. The best approach to this would be live interactive conversations in controlled experiments between study participants and our talking, gesturing AI ECA. Still, as seen in Study 1, current issues around safety and lockdown during COVID made live studies not possible. It should be noted that the gendered and non-gendered characters that I designed and validated in study 2 and 3 were made to work (in fully rigged emotional facial and character animated mode) in our advanced interactive ECA system, and I have programmed into the code features and upgrades to our ECA system that can better incorporate them. So, as you will see in future work below, I am ready when possible to begin interactive studies as future work.

Continuing from current research on the technical side, I would like to fully incorporate advanced AI ECA interactivity into the studies. Our current lab system is capable of this, including with empathy modelling; however, it was deemed essential to start with non-interactivity first, to control variables within this new line of research and, with that data, move on to face to face conversational interactivity. Interactivity with our ECAs implies using real-time and dynamic gestures, gaze, facial expressions and voice to the agents to create live interaction between agent and human, in a more advanced level, including a dialogue system to create embodied conversational agents which would be able to perform real-time face to face conversation with a partner verbally back and forth. Another reason I chose non-interactivity for this thesis work has to do with the real-time speech I used for a full face to face conversations and androgyny. That is, finding a unified voice for an androgynous agent is critical since no research-grade product is available to fit the androgynous character. Some experiments were conducted to alter our high-grade male and female character voices. Still, the early results with each were not at the level I deemed reasonable for an authentic androgynous voice, i.e. a voice where users would not assume a male or female gender. Therefore, the first future research step will be to find a unified voice

for an androgynous agent and measure if physical appearance with a voice can reinforce gender stereotypes.

For future research, I will replicate study 1 with the male, female and androgynous character to evaluate the stereotype assumption, eliminating the drawbacks of study 1 to provide fair and strong results. Additionally, I would like to replicate study 3, evaluating the similar hypothesis with fully interactive agents to see if that makes any difference than the static agent. The research's ultimate goal is not only to develop an effective aware embodied conversational agent that can perform real-time face-to-face conversation with a human interaction partner just like a human to human exchange but also that does not possess any kind of stereotype gender bias/attribution to the user experience. Thus, the ECAs can be effectively used in any role, such as health counsellor.

The research findings will be able to open a new dimension to the stereotypical perspective towards the gender of agents in virtual agent research. My hope is the research results from this and ongoing studies in this ECA and gender field educates researchers and companies to think beyond applications where female characters are being only used and instead rethink gender and gender neutral roles in ECAs of the future.

6.4   Final Words

Our society is changing its notions of gender stereotypical traits and roles. Virtual agent research should follow, where the old assumptions and research need to be revisited in light of new evidence on the critical gap in gender stereotypes. This research is the first step to highlight that these changing gender stereotypes also apply to virtual agents. By careful examination of the current stereotypical traits and introducing the androgynous agents' concept, I aimed to find an approach to reduce the stereotypical assumptions used by the virtual agent research, rather than reinforcing the bias. These findings showed that in designing representative male, female, androgynous characters with systematical manipulation of static facial cues. These results also mostly supported the findings of a gender-balanced view for stereotypical traits and roles from recent research on humans. I hope these findings would encourage constructive discussions in the research community to understand and

change our society's biases. I present the findings in the hope to foster discussions in virtual agent research and the frequent stereotypical use of gender representations. As there are very few studies on androgynous agents, I also hope that these design methodologies and results would encourage future work and further research in inclusive technologies. For future work, I would like to extend the evaluation in an interactive scenario, while controlling for dynamic variables such as gestures, gaze and expressions.

Bibliography

Andrea E Abele and Bogdan Wojciszke. Agency and communion from the perspective of self versus others. *Journal of personality and social psychology*, 93(5):751, 2007.

Andrea E. Abele, Nicole Hauke, Kim Peters, Eva Louvet, Aleksandra Szymkow, and Yanping Duan. Facets of the fundamental content dimensions: Agency with competence and assertiveness—communion with warmth and morality. *Frontiers in Psychology*, 7:1810, 2016. ISSN 1664-1078. doi: 10.3389/fpsyg.2016.01810.

Adobe. Adobe fuse cc (beta), 2017. URL `https://www.adobe.com/ca/products/fuse.html,(2015)`.

Elisabeth André and Catherine Pelachaud. Interacting with embodied conversational agents. In *Speech Technology: Theory and Applications*, pages 123–149. Springer US, Boston, MA, 2010. ISBN 9780387738185. doi: https://doi.org/10.1007/978-0-387-73819-2_8.

Michael Argyle. *Bodily Communication*, volume 5. January 1988. doi: 10.2307/2063160.

Ali Arya and Steve DiPaola. Face modeling and animation language for mpeg-4 xmt framework. *IEEE Transactions on Multimedia*, 9(6):1137–1146, 2007.

Ali Arya, Steve DiPaola, and Avi Parush. Perceptually valid facial expressions for character-based applications. *International Journal of Computer Games Technology*, 2009, 2009.

Robert K. Atkinson, Richard E. Mayer, and Mary Margaret Merrill. Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, 30(1):117–139, January 2005. ISSN 0361476X. doi: 10.1016/j.cedpsych.2004.07.001.

David Bakan. *The duality of human existence ; an essay on psychology and religion.* Rand McNally, 1966. doi: https://doi.org/10.1086/224499.

Susan A. Basow. *Gender: Stereotypes and Roles.* Thomson Brooks/Cole Publishing Co, 1992.

Amy Baylor and Suzanne J. Ebbers. Evidence that multiple agents facilitate greater learning. *Artificial intelligence in education: Shaping the future of learning through intelligent technologies*, pages 377–379, 2003.

Amy Baylor and Yanghee Kim. The role of gender and ethnicity in pedagogical agent perception. In *E-Learn: World Conference on E-Learning in Corporate, Government,*

Healthcare, and Higher Education, pages 1503–1506. Association for the Advancement of Computing in Education (AACE), 2003.

Amy Baylor, Jeeheon Ryu, and E. Shen. The effects of pedagogical agent voice and animation on learning, motivation and perceived persona. In *EdMedia+ Innovate Learning*, pages 452–458. Association for the Advancement of Computing in Education (AACE), 2003a.

Amy Baylor, E Shen, and Xiaoxia Huang. Which pedagogical agent do learners choose? the effects of gender and ethnicity. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, pages 1507–1510. Association for the Advancement of Computing in Education (AACE), 2003b.

Amy L. Baylor. The impact of pedagogical agent image on affective outcomes. In *International Conference on Intelligent User Interfaces, San Diego, CA*, page 29, 2005.

Amy L Baylor. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions Of The Royal Society B-Biological Sciences*, 364(1535):3559–3565, 2009. ISSN 0962-8436. doi: https://doi.org/10.1098/rstb.2009.0148.

Amy L. Baylor and Yanghee Kim. Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role. In James C. Lester, Rosa Maria Vicari, and Fábio Paraguaçu, editors, *Intelligent Tutoring Systems*, Lecture Notes in Computer Science, pages 592–603, Berlin, Heidelberg, 2004. Springer. ISBN 978-3-540-30139-4. doi: 10.1007/978-3-540-30139-4_56.

Ulysses Bernardet and Steve DiPaola. Affective response patterns as indicators of personality in virtual characters. In *CogSci*, 2014.

Ulysses Bernardet and Steve DiPaola. A framework for exogenous and endogenous reflexive behavior in virtual characters. In *International Conference on Intelligent Virtual Agents*, pages 132–136. Springer, 2015.

Ulysses Bernardet, Sin-hwa Kang, Andrew Feng, Steve DiPaola, and Ari Shapiro. A dynamic speech breathing system for virtual characters. In *International Conference on Intelligent Virtual Agents*, pages 43–52. Springer, 2017.

Sylvia Beyer. Low awareness of occupational segregation and the gender pay gap: No changes over a 16-year span. *Current Psychology*, 37(1):373–389, 2018. ISSN 1046-1310. doi: https://doi.org/10.1007/s12144-016-9521-4.

Timothy Bickmore and Justine Cassell. Social dialongue with embodied conversational agents. In *Advances in Natural Multimodal Dialogue Systems*, pages 23–54. Springer, 2005.

Timothy Bickmore, Everlyne Kimani, Ha Trinh, Alexandra Pusateri, Michael Paasche-Orlow, and Jared Magnani. Managing chronic conditions with a smartphone-based conversational virtual agent. In *Proceedings of the 18th International Conference on intelligent virtual agents*, IVA '18, pages 119–124. ACM, 2018. ISBN 9781450360135. doi: https://doi.org/10.1145/3267851.3267908.

Timothy W. Bickmore, Laura M. Pfeifer, Donna Byron, Shaula Forsythe, Lori E. Henault, Brian W. Jack, Rebecca Silliman, and Michael K. Paasche-Orlow. Usability of conversational agents by patients with inadequate health literacy: Evidence from two clinical trials. *Journal of health communication*, 15(S2):197–210, 2010a.

Timothy W. Bickmore, Kathryn Puskar, Elizabeth A. Schlenk, Laura M. Pfeifer, and Susan M. Sereika. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers*, 22(4):276–288, July 2010b. ISSN 1873-7951. doi: 10.1016/j.intcom.2010.02.001.

Timothy W Bickmore, Daniel Schulman, and Candace Sidner. Automated interventions for multiple health behaviors using conversational agents. *Patient education and counseling*, 92(2):142, 2013a. ISSN 0738-3991. doi: https://doi.org/10.1016/j.pec.2013.05.011.

Timothy W. Bickmore, Laura M. Pfeifer Vardoulakis, and Daniel Schulman. Tinker: A relational agent museum guide. *Autonomous Agents and Multi-Agent Systems*, 27(2): 254–276, September 2013b. ISSN 1573-7454. doi: 10.1007/s10458-012-9216-7.

Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124, 2002. ISSN 1047-840X. doi: https://doi.org/10.1207/S15327965PLI1302_03.

Sheryl Brahnam and Antonella De Angeli. Gender affordances of conversational agents. *Interacting With Computers*, 24(3):139–153, 2012. ISSN 0953-5438. doi: https://doi.org/10.1016/j.intcom.2012.05.001.

Adina Bresge. Why do popular digital assistants default to female voices?, mar 2019. URL https://www.ctvnews.ca/sci-tech/why-do-popular-digital-assistants-default-to-female-voices-1.4349613.

Elizabeth Brown and David I. Perrett. What gives a face its gender? *Perception*, 22(7): 829–840, 1993. doi: 10.1068/p220829.

Diana Burgess and Eugene Borgida. Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, public policy, and law*, 5(3):665, 1999.

D. Michael Burt and David I. Perrett. Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information. *Proceedings of the Royal Society B: Biological Sciences*, 259(1355):137–143, 1995. ISSN 0962-8452. doi: https://doi.org/10.1098/rspb.1995.0021.

Judith Butler. *Gender trouble: Feminism and the subversion of identity.* routledge, 2011.

J. E Cameron. Social identification and gender-related ideology in women and men. *British Journal of Social Psychology*, 40(1):59–77, 2001.

Rae Carlson. Sex differences in ego functioning: Exploratory studies of agency and communion. *Journal of consulting and clinical psychology*, 37(2):267–277, 1971. ISSN 0022-006X.

Antonin Carrier, Eva Louvet, Bruno Chauvin, and Odile Rohmer. The primacy of agency over competence in status perception. *Social Psychology*, 2014.

Justin Cassell. More Than Just Another Pretty Face: Embodied Conversational Agents. *Communications of the ACM*, 43(4), 2001.

Justine Cassell, Joseph Sullivan, Elizabeth Churchill, and Scott Prevost. *Embodied Conversational Agents*. MIT Press, 2000. ISBN 978-0-262-03278-0.

Susana Castillo, Philipp Hahn, Katharina Legde, and Douglas W. Cunningham. Personality Analysis of Embodied Conversational Agents. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pages 227–232, 2018.

Shelly Chaiken. Communicator physical attractiveness and persuasion. *Journal of Personality and social Psychology*, 37(8):1387, 1979.

Andrew J. Cowell and Kay M. Stanney. Embodiment and Interaction Guidelines for Designing Credible, Trustworthy Embodied Conversational Agents. In Thomas Rist, Ruth S. Aylett, Daniel Ballin, and Jeff Rickel, editors, *Intelligent Virtual Agents*, Lecture Notes in Computer Science, pages 301–309, Berlin, Heidelberg, 2003. Springer. ISBN 978-3-540-39396-2. doi: 10.1007/978-3-540-39396-2_50.

Doris M. Dehn and Susanne Van Mulken. The impact of animated interface agents: A review of empirical research. *International journal of human-computer studies*, 52(1): 1–22, 2000.

Manfred Diehl, Stephanie K Owen, and Lise M Youngblade. Agency and communion attributes in adults' spontaneous self-representations. *International journal of behavioral development*, 28(1):1–15, 2004.

Steve DiPaola. Extending the range of facial types. 1991.

Steve DiPaola. Intelligent expression-based character agent systems. In *International Workshop on Intelligent Virtual Agents*, pages 3–4. Springer, 2009.

Steve DiPaola. Face, portrait, mask: using a parameterised system to explore synthetic face space. In *Electronic Visualisation in Arts and Culture*, pages 213–227. Springer, 2013.

Steve DiPaola and Özge Nilay Yalçin. A multi-layer artificial intelligence and sensing based affective conversational embodied agent. In *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 91–92. IEEE, 2019.

Alice H. Eagly and Steven J. Karau. Role congruity theory of prejudice toward female leaders. *Psychological review*, 109(3):573, 2002.

Alice H. Eagly and Wendy Wood. *Social Role Theory*, page 458–476. SAGE Publications Ltd, 2012. doi: 10.4135/9781446249222.

Alice H. Eagly, Christa Nater, David I. Miller, Michele Kaufmann, and Sabine Sczesny. Gender stereotypes have changed: A cross-temporal meta-analysis of u.s. public opinion polls from 1946 to 2018. *American Psychologist*, 75(3):301–315, Apr 2020. ISSN 1935-990X, 0003-066X. doi: 10.1037/amp0000494.

Donald H Enlow and Mark G Hans. *Essentials of facial growth*. WB Saunders Company, 1996.

Alan Feingold. Gender differences in personality: A meta-analysis. *Psychological bulletin*, 116(3):429, 1994.

Susan T Fiske. Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2):67–73, 2018.

Jodi Forlizzi, John Zimmerman, Vince Mancuso, and Sonya Kwak. How interface agents affect interaction between humans and computers. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces*, pages 209–221, 2007.

Mary Ellen Foster. Enhancing human-computer interaction with embodied conversational agents. In *International Conference on Universal Access in Human-Computer Interaction*, pages 828–837. Springer, 2007.

Brandon Griggs. Why computer voices are mostly female - CNN. https://www.cnn.com/2011/10/21/tech/innovation/female-computer-voices, oct 2011.

Rosanna Guadagno, Kimberly Swinth, and Jim Blascovich. Social evaluations of embodied agents and avatars. *Computers in Human Behavior*, 27(6):2380–2385, 2011a. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2011.07.017. URL `http://search.proquest.com/docview/907928699/`.

Rosanna E. Guadagno, Kimberly R. Swinth, and Jim Blascovich. Social evaluations of embodied agents and avatars. *Computers in Human Behavior*, 27(6):2380–2385, 2011b.

Agneta Gulz, Felix Ahlner, and Magnus Haake. Visual femininity and masculinity in synthetic characters and patterns of affect. In *Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007 Proceedings*, volume 4738 of *Lecture Notes in Computer Science*, pages 654–665. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 9783540748885. doi: https://doi.org/10.1007/978-3-540-74889-2_57.

Magnus Haake and Agneta Gulz. Visual stereotypes and virtual pedagogical agents. *Journal of Educational Technology & Society*, 11(4):1–15, 2008.

Amy G Halberstadt and Martha B Saitta. Gender, nonverbal behavior, and perceived dominance: A test of the theory. *Journal of Personality and Social Psychology*, 53(2):257, 1987.

David L. Hamilton and Tina K. Trolier. Stereotypes and stereotyping: An overview of the cognitive approach. 1986.

Madeline E. Heilman. Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of social issues*, 57(4):657–674, 2001.

Hung-Hsuan Huang. Embodied Conversational Agents. In *The Wiley Handbook of Human Computer Interaction*, chapter 26, pages 599–614. John Wiley & Sons, Ltd, 2017. ISBN 978-1-118-97600-5. doi: 10.1002/9781118976005.ch26.

IBM. Watson Assistant | IBM Cloud. https://www.ibm.com/cloud/watson-assistant/, November 2018.

Katherine Isbister and Patrick Doyle. Design and evaluation of embodied conversational agents: A proposed taxonomy. In *The First International Joint Conference on Autonomous Agents & Multi-Agent Systems*, 2002.

Linda A. Jackson. *Physical Appearance and Gender: Sociobiological and Sociocultural Perspectives*. SUNY Press, January 1992. ISBN 978-0-7914-0823-0.

Nahid Karimaghalou, Ulysses Bernardet, and Steve DiPaola. A model for social spatial behavior in virtual characters. *Computer Animation and Virtual Worlds*, 25(3-4):505–517, 2014.

Rabia Khan and Antonella De Angeli. The attractiveness stereotype in the evaluation of embodied conversational agents. In *IFIP Conference on Human-Computer Interaction*, pages 85–97. Springer, 2009.

Yanghee Kim. The Role of Agent Age and Gender for Middle-Grade Girls. *Computers in the Schools*, 33(2):59–70, 2016. ISSN 0738-0569. doi: 10.1080/07380569.2016.1143753.

Yanghee Kim, Amy L. Baylor, and PALS Group. Pedagogical agents as learning companions: The role of agent competency and type of interaction. *Educational Technology Research and Development*, 54(3):223–243, 2006.

Yanghee Kim, A. L Baylor, and E Shen. Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of computer assisted learning*, 23(3):220, 2007. ISSN 0266-4909. doi: https://doi.org/10.1111/j.1365-2729.2006.00210.x.

Tomoko Koda and Pattie Maes. Agents with faces: The effect of personification. In *Proceedings 5th IEEE International Workshop on Robot and Human Communication. RO-MAN'96 TSUKUBA*, pages 189–194, November 1996. doi: 10.1109/ROMAN.1996.568812.

Tomoko Koda and Yuuki Takeda. Perception of Culture-specific Gaze Behaviors of Agents and Gender Effects. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, pages 138–143, 2018.

Stefan Kopp, Lars Gesellensetter, Nicole C Krämer, and Ipke Wachsmuth. A conversational agent as museum guide – design and evaluation of a real-world application. In *Intelligent Virtual Agents*, volume 3661 of *Lecture Notes in Computer Science*, pages 329–343. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005. ISBN 3540287388.

Philipp Kulms, Nicole C. Krämer, Jonathan Gratch, and Sin-Hwa Kang. It's in their eyes: A study on female and male virtual humans' gaze. In Hannes Högni Vilhjálmsson, Stefan Kopp, Stacy Marsella, and Kristinn R. Thórisson, editors, *Intelligent Virtual Agents*, Lecture Notes in Computer Science, page 80–92. Springer, 2011. ISBN 978-3-642-23974-8. doi: 10.1007/978-3-642-23974-8_9.

Adrienne LaFrance. Why Do So Many Digital Assistants Have Feminine Names? https://www.theatlantic.com/technology/archive/2016/03/why-do-so-many-digital-assistants-have-feminine-names/475884/, March 2016.

Marianne LaFrance. Gender gestures: Sex, sex-role, and nonverbal communication. In *Gender and nonverbal behavior*, pages 129–150. Springer, 1981.

Michael Lankes, Regina Bernhaupt, and Manfred Tscheligi. An experimental setting to measure contextual perception of embodied conversational agents. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, pages 56–59, 2007.

Brenda Laurel and S. Joy Mountford. *The Art of Human-Computer Interface Design*. Addison-Wesley Longman Publishing Co., Inc., USA, 1990. ISBN 978-0-201-51797-2.

Max Louwerse, Arthur Graesser, Shulan Lu, and Heather Mitchell. Social cues in animated conversational agents. *Applied Cognitive Psychology*, 19(6):693–704, 2005. ISSN 08884080. doi: https://doi.org/10.1002/acp.1117. URL `http://search.proquest.com/docview/211292639/`.

Carol Lynn Martin and Diane N. Ruble. Patterns of gender development. *Annual review of psychology*, 61:353–381, 2010. ISSN 0066-4308. doi: 10.1146/annurev.psych.093008.100511.

Dominic W. Massaro. Symbiotic value of an embodied agent in language learning. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of The*, pages 10–pp. IEEE, 2004.

Dominic W Massaro, Michael M Cohen, Sharon Daniel, and Ronald A Cole. Chapter 7 - developing and evaluating conversational agents. In *Human Performance and Ergonomics*, pages 173–194. Academic Press, 1999. ISBN 9780123227355. doi: https://doi.org/10.1016/B978-012322735-5/50008-7.

Vincent Maya, Myriam Lamolle, and Catherine Pelachaud. Influences and embodied conversational agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1306–1307. IEEE Computer Society, 2004.

Kristen N Moreno, Natalie K Person, Amy B Adcock, RNV Eck, G Tanner Jackson, and Johanna C Marineau. Etiquette and efficacy in animated pedagogical agents: The role of stereotypes. In *AAAI symposium on personalized agents, Cape Cod, MA*, page 4, 2002.

Hazel Morton, Helen McBreen, and Mervyn Jack. Experimental Evaluation of the Use of ECAs in ecommerce Applications. In Zsófia Ruttkay and Catherine Pelachaud, editors, *From Brows to Trust: Evaluating Embodied Conversational Agents*, Human-Computer Interaction Series, pages 293–321. Springer Netherlands, Dordrecht, 2004. ISBN 978-1-4020-2730-7. doi: 10.1007/1-4020-2730-3_11.

Procheta Nag and Özge Nilay Yalçın. Effect of gender on interaction with embodied conversational agents. In *Proceedings of the 14th Women in Machine Learning Workshop co-located with NeurIPS'19 conference*, WIML '19, page 1, 2019. doi: https://doi.org/10.13140/RG.2.2.21434.62400.

Procheta Nag and Özge Nilay Yalçın. Gender stereotypes in virtual agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, IVA '20, New

York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450375863. doi: 10.1145/3383652.3423876.

Clifford Nass, Youngme Moon, and Nancy Green. Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices. *Journal of Applied Social Psychology*, 27(10):864–876, 1997. ISSN 1559-1816. doi: 10.1111/j.1559-1816.1997.tb00275.x.

Andreea Niculescu, Frans Van Der Sluis, and Anton Nijholt. Feminity, masculinity and androgyny: how humans perceive the gender of anthropomorphic agents. In *Proceedings of 13th International Conference on Human-Computer Interaction, Springer Verlag, Heidelberg*, pages 628–632, 2009.

Michael Nixon, Steve DiPaola, and Ulysses Bernardet. An eye gaze model for controlling the display of social status in believable virtual humans. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 1–8. IEEE, 2018.

Nikolaas N Oosterhof and Alexander Todorov. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32):11087–11092, 2008.

Alice O'Toole, Kenneth Deffenbacher, Dominique Valentin, Karen McKee, David Huff, and Hervé Abdi. The perception of face gender: The role of stimulus structure in recognition and classification. *Memory & Cognition*, 26(1):146–160, 1998. ISSN 0090-502X. doi: https://doi.org/10.3758/BF03211378.

PANDORABOTS. Mitsuku. https://www.pandorabots.com/mitsuku/, 2013.

Jeunese Payne, Andrea Szymkowiak, Paul Robertson, and Graham Johnson. Gendering the machine: Preferred virtual assistant gender and realism in self-service. In *International Workshop on Intelligent Virtual Agents*, pages 106–115. Springer, 2013.

Catherine Pelachaud. Multimodal expressive embodied conversational agents. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, pages 683–689, 2005.

Catherine Pelachaud. Greta: An interactive expressive embodied conversational agent. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 5–5, 2015.

Byron Reeves. *The media equation: How people treat computers, television, and new media like real people and places.* CSLI Publications ; Cambridge University Press, 1998. ISBN 1575860538.

Lazlo Ring, Timothy Bickmore, and Paola Pedrelli. An affectively aware virtual therapist for depression counseling. In *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) Workshop on Computing and Mental Health*, pages 01951–12, 2016.

Tony Roberts and Vicki Bruce. Feature saliency in judging the sex and familiarity of faces. *Perception*, 17(4):475–481, 1988. ISSN 0301-0066.

Rb Rosenberg-Kima, AL Baylor, Ea Plant, and CE Doerr. Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers In Human Behavior*, 24(6):2741–2756, 2008. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2008.03.017.

Zsófia Ruttkay, Claire Dormann, and Han Noot. Embodied conversational agents on a common ground. In *From Brows to Trust: Evaluating Embodied Conversational Agents*, volume 7 of *Human-Computer Interaction Series*, pages 27–66. Springer Netherlands, Dordrecht, 2004. ISBN 9781402027291. doi: https://doi.org/10.1007/1-4020-2730-3_2.

Maryam Saberi, Ulysses Bernardet, and Steve DiPaola. An architecture for personality-based, nonverbal behavior in affective virtual humanoid character. *Procedia Computer Science*, 41:204–211, 2014.

Noah L. Schroeder and Olusola O. Adesope. Impacts of pedagogical agent gender in an accessible learning environment. *Journal of Educational Technology & Society*, 18(4): 401–411, 2015.

Daniel Schulman and Timothy Bickmore. Persuading users through counseling dialogue with a conversational agent. In *Proceedings of the 4th International Conference on Persuasive Technology*, pages 1–8, 2009.

Daniel Schulman, Timothy Bickmore, and Candace Sidner. An intelligent conversational agent for promoting long-term health behavior change using motivational interviewing. In *2011 AAAI Spring Symposium Series*, 2011.

Eric Hal Schwartz. New Zealand Police Enlist Virtual AI Officer. https://voicebot.ai/2020/02/20/new-zealand-police-enlist-virtual-ai-officer/, February 2020.

Sabine Sczesny, Christa Nater, Alice H. Eagly, Christa Nater, and Alice H. Eagly. Agency and communion: Their implications for gender stereotypes and gender identities, Sep 2018. URL http://www.taylorfrancis.com/.

Ameneh Shamekhi, Q Liao, Dakuo Wang, Rachel Bellamy, and Thomas Erickson. Face value? exploring the effects of embodiment for a group facilitation agent. In *Proceedings of the 2018 CHI Conference on human factors in computing systems*, volume 2018- of *CHI '18*, pages 1–13. ACM, 2018. ISBN 9781450356206. doi: https://doi.org/10.1145/3173574.3173965.

Annika Silvervarg. How students perceive the gender and personality of a visually androgynous agent. In *International Conference on Intelligent Virtual Agents*, pages 420–423. Springer, 2016.

Annika Silvervarg and Arne Jönsson. Subjective and objective evaluation of conversational agents in learning environments for young teenagers. In *Proceedings of the 7th IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, 2011.

Annika Silvervarg, Kristin Raukola, Magnus Haake, and Agneta Gulz. The effect of visual gender on abuse in conversation with ecas. In *Intelligent Virtual Agents: 12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012. Proceedings*, volume 7502 of *Lecture Notes in Computer Science*, pages 153–160. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. ISBN 9783642331961. doi: https://doi.org/10.1007/978-3-642-33197-8_16.

Annika Silvervarg, Magnus Haake, and Agneta Gulz. Educational potentials in visually androgynous pedagogical agents. In *International Conference on Artificial Intelligence in Education*, pages 599–602. Springer, 2013.

Megan Specia. Siri and alexa reinforce gender bias, u.n. finds, may 2019. URL `https://www.nytimes.com/2019/05/22/world/siri-alexa-ai-gender-bias.html`.

United States. Bureau Of Labor Statistics. *Women in the labor force: a databook*. Number 1084. Dec 2019. URL `https://www.bls.gov/opub/reports/womens-databook/2019/pdf/home.pdf`.

U.S. BUREAU OF LABOR STATISTICS. Median weekly earnings of full-time wage and salary workers by detailed occupation and sex. https://www.bls.gov/cps/cpsaat39.htm, jan 2020.

Wolfgang Stroebe, Chester A. Insko, Vaida D. Thompson, and Bruce D. Layton. Effects of physical attractiveness, attitude similarity, and sex on various aspects of interpersonal attraction. *Journal of Personality and Social Psychology*, 18(1):79, 1971.

Clare A M Sutherland, Andrew W Young, Carmen A Mootz, and Julian A Oldmeadow. Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British journal of psychology (London, England : 1953)*, 106(2):186, 2015. ISSN 0007-1269. doi: https://doi.org/10.1111/bjop.12085.

Marcus Thiebaux, Stacy Marsella, Andrew N Marshall, and Marcelo Kallmann. Smartbody: Behavior realization for embodied conversational agents. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pages 151–158, 2008.

Ha Trinh, Ameneh Shamekhi, Everlyne Kimani, and Timothy Bickmore. Predicting user engagement in longitudinal interventions with virtual agents. In *Proceedings of the 18th International Conference on intelligent virtual agents*, IVA '18, pages 9–16. ACM, 2018. ISBN 9781450360135. doi: https://doi.org/10.1145/3267851.3267909.

Christiana Tsiourti, Emilie Joly, Cindy Wings, Maher Ben Moussa, and Katarzyna Wac. Virtual assistive companions for older adults: Qualitative field study and design implications. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. ICST, 2014. ISBN 978-1-63190-011-2. doi: 10.4108/icst.pervasivehealth.2014.254943. URL `http://eudl.eu/doi/10.4108/icst.pervasivehealth.2014.254943`.

Jeremy Owen Turner, Michael Nixon, Ulysses Bernardet, and Steve DiPaola. *Integrating cognitive architectures into virtual character design*. IGI Global, 2016.

Kazuya Ueki, Hiromitsu Komatsu, Satoshi Imaizumi, Kenichi Kaneko, Nobuhiro Sekine, Jiro Katto, and Tetsunori Kobayashi. A method of gender classification by integrating facial, hairstyle, and clothing images. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 4, pages 446–449. IEEE, 2004.

Dina Utami, Timothy Bickmore, Asimina Nikolopoulou, and Michael Paasche-Orlow. Talk About Death: End of Life Planning with a Virtual Agent. In Jonas Beskow, Christopher Peters, Ginevra Castellano, Carol O'Sullivan, Iolanda Leite, and Stefan Kopp,

editors, *Intelligent Virtual Agents*, Lecture Notes in Computer Science, pages 441–450, Cham, 2017. Springer International Publishing. ISBN 978-3-319-67401-8. doi: 10.1007/978-3-319-67401-8_55.

Marco Vala, Gabriel Blanco, and Ana Paiva. Providing Gender to Embodied Conversational Agents. In Hannes Högni Vilhjálmsson, Stefan Kopp, Stacy Marsella, and Kristinn R. Thórisson, editors, *Intelligent Virtual Agents*, Lecture Notes in Computer Science, pages 148–154, Berlin, Heidelberg, 2011. Springer. ISBN 978-3-642-23974-8. doi: 10.1007/978-3-642-23974-8_16.

Anu Venkatesh, Chandra Khatri, Ashwin Ram, Fenfei Guo, Raefer Gabriel, Ashish Nagar, Rohit Prasad, Ming Cheng, Behnam Hedayatnia, Angeliki Metallinou, et al. On evaluating and comparing conversational agents. *arXiv preprint arXiv:1801.03625*, 4:60–68, 2018.

Janet H Walker, Lee Sproull, and R Subramani. Using a human face in an interface. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 85–91, 1994.

Catharine Wang, Timothy Bickmore, Deborah J. Bowen, Tricia Norkunas, MaryAnn Campion, Howard Cabral, Michael Winter, and Michael Paasche-Orlow. Acceptability and feasibility of a virtual counselor (VICKY) to collect family health histories. *Genetics in Medicine*, 17(10):822–830, October 2015. ISSN 1530-0366. doi: 10.1038/gim.2014.198.

Mark West, Rebecca Kraut, and Han Ei Chew. I'd blush if i could: closing gender divides in digital skills through education. 2019.

Preben Wik and Anna Hjalmarsson. Embodied conversational agents in computer assisted language learning. *Speech Communication*, 51(10):1024–1037, 2009. ISSN 0167-6393. doi: https://doi.org/10.1016/j.specom.2009.05.006. URL `http://search.proquest.com/docview/85705901/`.

John E. Williams and Deborah L. Best. *Measuring Sex Stereotypes: A Multination Study, Rev.* Sage Publications, Inc, 1990.

Özge Nilay Yalçin and Steve DiPaola. A computational model of empathy for interactive agents. *Biologically inspired cognitive architectures*, 26:20–25, 2018.

Ozge Nilay Yalçın and Steve DiPaola. Evaluating levels of emotional contagion with an embodied conversational agent. In *Proceedings of the 41st annual conference of the cognitive science society*, 2019.

Masami K Yamaguchi, Tastu Hirukawa, and So Kanazawa. Judgment of gender through facial parts. *Perception*, 42(11):1253–1265, 2013. ISSN 0301-0066. doi: https://doi.org/10.1068/p240563n.

Leslie Zebrowitz. *Reading faces: Window to the soul?* Routledge, 2018.

Shuo Zhou, Timothy Bickmore, Amy Rubin, Catherine Yeksigian, Molly Sawdy, and Steven R. Simon. User Gaze Behavior while Discussing Substance Use with a Virtual Agent. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pages 353–354, 2018.

John Zimmerman, Ellen Ayoob, Jodi Forlizzi, and Mick McQuaid. Putting a Face on Embodied Interface Agents. January 2005. doi: 10.1184/R1/6470366.v1.

Appendix A

Ethics Approval

Amendment Approval - Delegated

Study Number: 2019s0520
Study Title: Designing Conversational Agents for Interactions Across the Lifespan

Amendment Approval Date: September 16, 2020   **Expiry Date:** January 23, 2021
Principal Investigator: Moreno, Sylvain   **Supervisor:** N/A
SFU Position: Faculty   **Faculty/Department:** SIAT

SFU Collaborator: DiPaola, Steve
External Collaborator: N/A
Research Personnel: Nag, Procheta
Project Leader: N/A

Funding Source: AGE-WELL
Funding Title: National Innovation Hub: Digital Media and Technologies in Health

Document(s) Approved in this Amendment:
- ORE Amendment Request Form, uploaded 2020, September 16
- Change of PI Form, uploaded 2020, September 16
- Study Detail, version 3 dated 2020, September 15
- Consent Form AMT, version 2 dated 2020, September 15
- Consent Form, version 3 dated 2020, September 15

The amendment(s) for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human participants.

Please note that approval of the amendment(s) does not change the expiry date on the current SFU REB approval of this study. The approval for this study expires on the **Expiry Date. An annual renewal form must be completed every year prior to the Expiry Date. Failure to submit an annual renewal form will lead to your study being suspended and potentially terminated.**

This letter is your official Amendment Approval documentation for this project. Please keep this document for reference purposes.

The amendment to this study has been approved by an authorized delegated reviewer.

Appendix B

Study I process

Effect of gender with embodied conversational agent interaction

Effect of gender during the interaction with embodied conversational agent

In this study, you will watch two short video clips that will appear to be very similar, but do not assume that the conversational material is exact and instead please just give your natural response to what you've just seen. After each clip, you will get a short questionnaire. You agree not to disclose the content of the study to others.

* 2. Do you agree and consent to these terms?

◯ I agree to participate on the study.　　　◯ No thanks, I am not interested.

Effect of gender with embodied conversational agent interaction
Welcome to the Survey

You will see two video clips where a human person is talking with a embodied conversational agent. The avatar has facial, emotional and body gestures reflecting human like behavior.

Please give your natural response. Your feedback is important.

Just two quick information before starting the survey....................

* 3. What is your gender?

| Female | Male | Others |
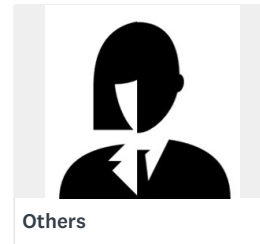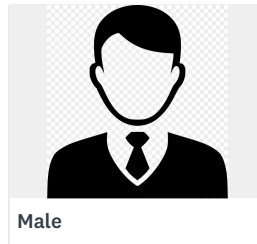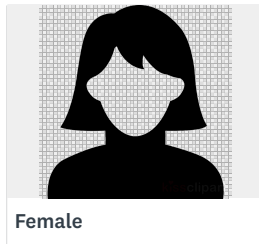
* 4. What is your age?

◯ Under 18   ◯ 18-25   ◯ 26-35   ◯ 36+

Effect of gender with embodied conversational agent interaction
Embodied conversational agent talking with a human person

Please watch this video till the end at full screen (e.g. clicking YouTube icon) with audio on. After watching this video you will be asked to give your opinion
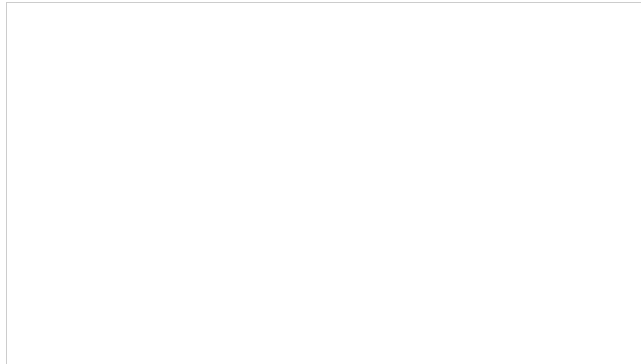
Effect of gender with embodied conversational agent interaction
Embodied conversational agent survey questionnaire
Please choose your option about how did you feel about watching the video

* 5. How comfortable would you be working with this embodied agent?

Very comfortable	Comfortable	Possibly	Uncomfortable	Very uncomfortable
○	○	○	○	○

* 6. How engaging the conversation seems to you?

Highly engaging	Engaging	Moderately	Unengaging	Highly unengaging
○	○	○	○	○

* 7. How interested would you be to have this embodied agent as your assistant?

Highly interested	Interested	Probably	Uninterested	Highly uninterested
○	○	○	○	○

* 8. How would you describe the personality and attitude of the embodied agent?

Very helpful	Helpful	Neutral	Not helpful	Very unhelpful
○	○	○	○	○

* 9. The embodied agent provides natural response through voice, facial and body expression

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
○	○	○	○	○

10. Please list a few words you would use to describe the personality and attitude of this embodied agent? Please write your short answer in the box below --

Effect of gender with embodied conversational agent interaction
Embodied conversational agent talking with a human person

Please watch this video till the end at full screen (e.g. clicking YouTube icon) with audio on. After watching this video you will be asked to give your opinion
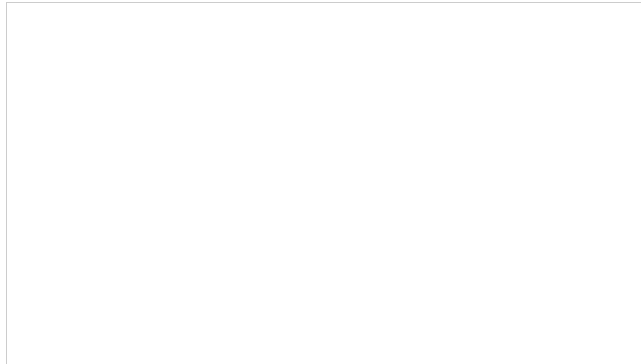
Effect of gender with embodied conversational agent interaction

Embodied conversational agent survey questionnaire

Please choose your option about how did you feel about watching the video

* 11. How comfortable would you be working with this embodied agent?

Very comfortable	Comfortable	Possibly	Uncomfortable	Very uncomfortable
○	○	○	○	○

* 12. How engaging the conversation seems to you?

Highly engaging	Engaging	Moderately	Unengaging	Highly unengaging
○	○	○	○	○

* 13. How interested would you be to have this embodied agent as your assistant?

Highly interested	Interested	Probably	Uninterested	Highly uninterested
○	○	○	○	○

* 14. How would you describe the personality and attitude of the embodied agent?

Very helpful	Helpful	Neutral	Not helpful	Very unhelpful
○	○	○	○	○

* 15. The embodied agent provides natural response through voice, facial and body expression

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
○	○	○	○	○

16. Please list a few words you would use to describe the personality and attitude of this embodied agent? Please write your short answer in the box below --

Appendix C

Study I participants comment

Partici-pants id	Comments for male ECA	Comments for female ECA
19	Too many stagnant pauses! It made everything feel bad.	once again, hella stilted, her eyes disturbed me.
22	strange, monotone	
25	The personality and attitude seems friendly.	The agent seems helpful and more engaged in the conversation.
26	its hard to have a human interaction with a robot	
27	The responses are appropriate, however the voice is too monotone and robotic.	The fidgeting of the assistant was a little distracting. The attitude and personality were good.
28	Response is too slow.	
29	Dull, Robotic, Creepy	Friendly, kind, less creepy
30	very robotic	very robotic still
31		soft and gentle
32	The embodied agent seems very disingenuous. The tone of his voice is flat and emotive making him difficult to relate to. I don't think I've seen him blink once, I felt more scared than calm.	The embodied agent struggles with similar issues as the first. However the fact that she moves around more and seems more emotive makes her seem more realistic and makes me more comfortable in comparison.
33	I would feel more engaged if I was actually speaking with the assistant rather then watching a video with such a time gap.	Much friendlier, better tone of voice then the male assistant.
34	awkward silences in between makes everything else feel off.	It feels better than the first one. I think it's because the voice doesn't sound as harsh.
37	friendly, trying to be helpful	better to show a bit more caring
40	Very dry	
41	robotic agent	
43	He is emotionless.	Speak without pace changes and eyeballs don't move.
45	Pretty normal with no surprised.	Not bad
46	I think he's lack of emotion, use same facial expression to say different things.	Attitude is good within the various facial expressions.
47	Unnatural facial expression	Better than the last one

Table C.1: Open ended question responses

48	Boring, uninspiring, unnatural, robotic	Better than the male voice- but the intonation and tone of the embodied agent still suffers from a lack of enthusiasm or nuances found in natural speech. Personality is neutral and attitude is neutral which causes the agent to suffer from being uninspiring to have as an assistant. However, the higher pitch associated with female voices helps liven up boring conversations and makes it slightly more engaging.
49	The responding time takes too long	Still feels like a machine
51	i hope he can watch to my eyes all the time not look around make me feel not respect	
52	slow, robotic, stiff, rote, simplistic	slow, a bit twitchy, sometimes empathetic looking, simplistic, nice
54	slow to react with a monotone voice	
55	Tried to provide helpful advices but didn't sound sincere. The pauses before responses were awkwardly long.	The response to the first answer seemed off. The pauses before responses felt less long than the previous one.
56	Slow, distant, some responses seemed too certain despite a clear lack of information provided	Forward, assertive, somewhat slow to respond
57	slow with some random unnecessary moments such as fixing his hair	pleasant, kind but still slow and does some weird movements
58	empathetic kind disinterested non-genuine	empathetic kind disinterested placid
59	the agent is trying to be helpful, but the pauses are too long so it is awkward	personality is fine, but the conversation is unengaging due to long pauses.
61	empathetic, caring, thoughtful	professional, caring, distracted

Table C.2: Open ended question responses (continue...)

62	seems very neutral, not enough personality! seems professional, put together, somewhat empathetic but tone of voice is off	Kind, interested, her voice matches what she looks like
63	Off-putting long silences. Voice emphases are highly unnatural. Gaze is indecipherable. Autonomic actions distracting.	Virtually the same as the last. Unnatural voice synthesis, long pauses, difficult-to-interpret autonomic actions. Flat affect.
64	Flat, unemotional, contrived.	Canned response, stiff, generic
65	Robotic, unfeeling	Distracted, kind, but flat
66	Pausing to understand the speech is the most disengaging part	Professional
67	Attentive, professional	Slightly more engaging than the previous one
68	Helpful and empathatic. The expressions were appropriate, however, the long delays in avatar response were very disengaging.	Little work needed on rendering round the mouth. Avatar seemed helpful and empathatic
70	I'd say the avatar was trying to be helpful, but the long pauses definitely breaks the immersion. Additionally, the avatar seems to be smiling a bit too much in my opinion.	Same with the other avatar, it seems to try to be helpful but the long pauses are a bit uncomfortable. It also feels like her face doesn't change emotion too much so it's hard to see what the avatar is feeling.
71	The avatar's voice, visual style and animations make me feel extremely uncomfortable. Even the answers are generic and clearly not prescribed by a professional psychologist.	Same feel as the first one, but slightly less creepy - I'm not too sure why...
73	I would say the avatar is helpful, however the lack of emotion on its face and voice is off putting	I prefer this one before the other

Table C.3: Open ended question responses (end)

75	The avatar itself was good looking and its attitude was good. However, because the response time was not immidiate, I could feel big pauses inside the convrsation	The voice was really calm and soothing. Also, the expressions of the avatar were good. However, there has been a small delay for the answers. other than that great experience
77	The avatar seems concerned and engaged. Although, this is true, I still don't feel so comfortable having an avatar be my assistant!	The avatar seemed quite concerned although some gestures seemed a bit off (like when she looked down to the right side). She seemed more approachable than the male avatar though.
78	Unphased, cold, robotic, slow	Nice, concerned
80	the answers from the avatar caused the avatar to felt like his personality is quite positive, though he gave some very detailed answers right after asking	helpful, quite positive, feels slightly more engaging possibly due to timing

Table C.4: Open ended question responses (end)

Appendix D

Study 2 process

D.1 First Iteration of Character Design

Visual Appearance of 3D Characters-study 1
Visual Appearance of 3D Characters-phase 1

In this study, you will see a series of portraits of 3D characters, and you will be asked to answer a questionnaire about your opinion about those images. The survey will take approximately 5 minutes.

We do not think there is anything in this study that could harm you or be bad for you. There are no foreseeable risks to you in participating in this study.

 * 2. Do you agree and consent to these terms?

○ I agree to participate on the study.   ○ No thanks, I am not interested.

Visual Appearance of 3D Characters-study 1
Welcome to the Survey
Just two quick questions before you start the survey.

* 3. What is your gender?
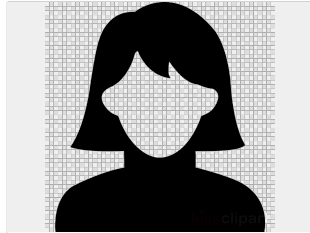
Female

Male
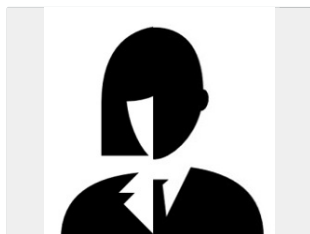
Others

* 4. What is your age?

◯ Under 18 ◯ 18 – 24 ◯ 25-34 ◯ 35-44 ◯ 45-54 ◯ 55-64 ◯ Above 65

* 5. Are you familiar with 3D characters?

◯ Yes ◯ No ◯ A little bit

* 6. Have you done any courses or worked on 3D modeling?

◯ Yes ◯ No

Visual Appearance of 3D Characters-study 1
Appearance of 3D characters
You will see several variations of the look of 3D humanoid characters. We want to know your thought of the character's appearance or how they look, which character you think as or more look like a female, male or androgynous.

*** Androgynous means someone who presents as not particularly female or male look. It carries some qualities from male and some qualities from the female appearance.

From next page, you will see some portraits of the humanoid 3D characters on each page and you have to answer the question. Please give your natural response. If you are ready to start the survey click "Next".

Visual Appearance of 3D Characters-study 1
3D Characters
See the images below carefully

* 7. What do you think of the gender of this 3D character?

Male	Androgynous	Female

8. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters-study 1

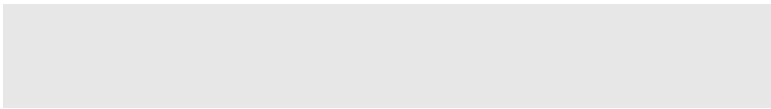* 9. What do you think of the gender of this 3D character?

Male	Androgynous	Female

10. If have any opinion about the character's design of gender please write down below.
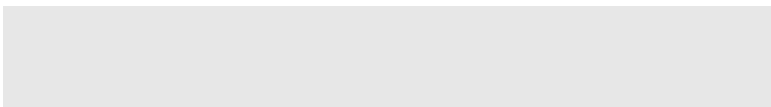
Visual Appearance of 3D Characters-study 1

* 11. What do you think of the gender of this 3D character?

| Male | Androgynous | Female |

12. If have any opinion about the character's design of gender please write down below.

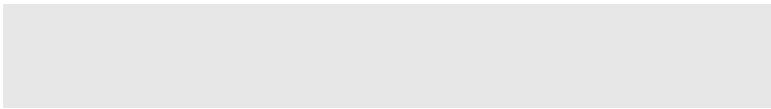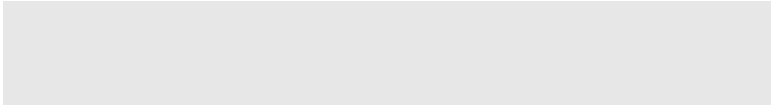# Visual Appearance of 3D Characters-study 1
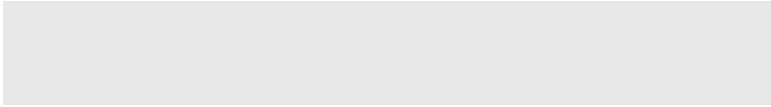

* 13. What do you think of the gender of this 3D character?

Male	Androgynous	Female

14. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters-study 1

* 15. What do you think of the gender of this 3D character?

Male	Androgynous	Female

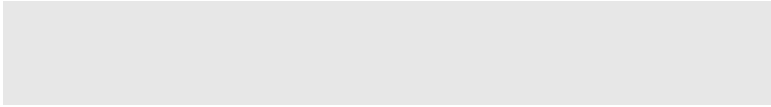16. If have any opinion about the character's design of gender please write down below.
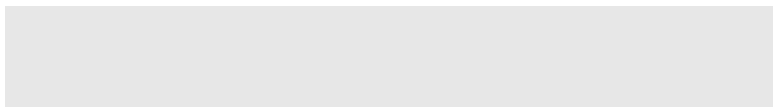
Visual Appearance of 3D Characters-study 1

* 17. What do you think of the gender of this 3D character?

Male	Androgynous	Female

18. If have any opinion about the character's design of gender please write down below.

D.2   Second Iteration of Character Design

Visual Appearance of 3D Characters- study 2
Visual Appearance of 3D characters-phase 2

In this study, you will see a series of portraits of 3D characters, and you will be asked to answer a questionnaire about your opinion about those images. The survey will take approximately 5 minutes.

We do not think there is anything in this study that could harm you or be bad for you. There are no foreseeable risks to you in participating in this study.

* 2. Do you agree and consent to these terms?

◯ I agree to participate on the study.　　◯ No thanks, I am not interested.

Visual Appearance of 3D Characters- study 2
Welcome to the Survey
Just two quick questions before you start the survey.

* 3. What is your gender?
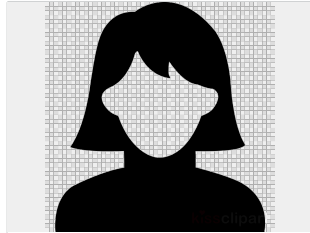
Female

Male
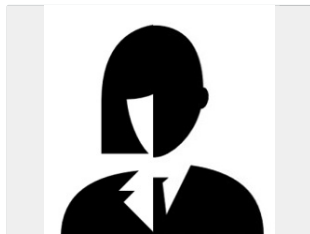
Others

* 4. What is your age?

◯ Under 18 ◯ 18 – 24 ◯ 25-34 ◯ 35-44 ◯ 45-54 ◯ 55-64 ◯ Above 65

* 5. Are you familiar with 3D characters?

◯ Yes ◯ No ◯ A little bit

* 6. Have you done any courses or worked on 3D modeling?

◯ Yes ◯ No

Visual Appearance of 3D Characters- study 2
Appearance of 3D characters
You will see several variations of the look of 3D humanoid characters. We want to know your thought of the character's appearance or how they look, which character you think as or more look like a female, male or androgynous.

*** Androgynous means someone who presents as not particularly female or male look. It carries some qualities from male and some qualities from the female appearance.

From next page, you will see some portraits of the humanoid 3D characters on each page and you have to answer the question. Please give your natural response. If you are ready to start the survey click "Next".
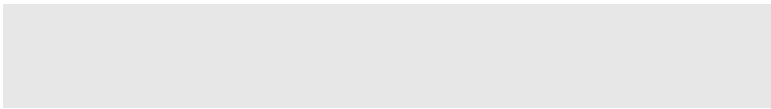
Visual Appearance of 3D Characters- study 2

* 7. What do you think of the gender of this 3D character?

Male	Androgynous	Female

8. If have any opinion about the character's design of gender please write down below.
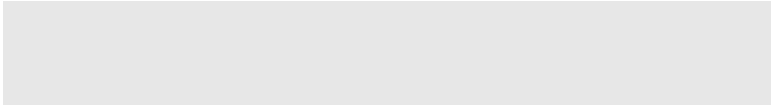
Visual Appearance of 3D Characters- study 2

* 9. What do you think of the gender of this 3D character?

Male	Androgynous	Female

10. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 11. What do you think of the gender of this 3D character?

Male	Androgynous	Female

12. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 13. What do you think of the gender of this 3D character?

Male	Androgynous	Female

14. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 15. What do you think of the gender of this 3D character?

Male	Androgynous	Female

16. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 17. What do you think of the gender of this 3D character?

Male	Androgynous	Female

18. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 19. What do you think of the gender of this 3D character?

Male	Androgynous	Female

20. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 21. What do you think of the gender of this 3D character?

Male	Androgynous	Female

22. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 23. What do you think of the gender of this 3D character?

Male	Androgynous	Female

24. If have any opinion about the character's design of gender please write down below.

Visual Appearance of 3D Characters- study 2

* 25. What do you think of the gender of this 3D character?

Male	Androgynous	Female

26. If have any opinion about the character's design of gender please write down below.

Appendix E

Study 3 process

Visual Appearance of Virtual Agents

In this study, you will see pictures of three virtual agents, and you will be asked to answer some questions about how do you feel about them. The survey will take approximately 5 minutes.

* 2. Do you agree and consent to these terms?

◯ **I agree to participate on the study.**   ◯ **No thanks, I am not interested.**

Demographic Questions

* 3. What is your age?

○ Under 18 ○ 18-24 ○ 25-34 ○ 35-44 ○ 45-54 ○ 55-64 ○ 65+

* 4. What is your gender?

○ Female

○ Male

○ Other (please specify)

5. Are you familiar with any of the systems?

☐ 3D character

☐ Virtual agent

☐ Voice Assistant

☐ Chatbot

Welcome to the Survey

Virtual Agent: A virtual agent is a computer-generated 3D animated, virtual character. They can be used to have a conversation with users with both verbal and non-verbal behavior like gestures and facial expressions.

Starting from the next page, you will see portraits of different virtual agents and will be asked about your opinions of them.

If you are ready to start the survey click "Next".

Embodied conversational agent survey questionnaire

* 6. In your opinion, which of these the following words defines the agent you see in the picture above?
The agent looks ...

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Affectionate	○	○	○	○	○
Ambitious	○	○	○	○	○
Aggressive	○	○	○	○	○
Creative	○	○	○	○	○
Courageous	○	○	○	○	○
Compassionate	○	○	○	○	○
Decisive	○	○	○	○	○
Emotional	○	○	○	○	○
Intelligent	○	○	○	○	○
Innovative	○	○	○	○	○
Organized	○	○	○	○	○
Sensitive	○	○	○	○	○

* 7. In your opinion, which one of the roles below would fit the agent you have seen in the picture above?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Personal Assistant	○	○	○	○	○
Health Counsellor	○	○	○	○	○
Exercise Coach	○	○	○	○	○
Financial Advisor	○	○	○	○	○
News Reporter	○	○	○	○	○
Storyteller	○	○	○	○	○

8. Please indicate why do you think you felt this way about this agent? This question is optional.

Embodied conversational agent survey questionnaire

* 9. In your opinion, which of these the following words defines the agent you see in the picture above?
The agent looks ...

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Affectionate	○	○	○	○	○
Ambitious	○	○	○	○	○
Aggressive	○	○	○	○	○
Creative	○	○	○	○	○
Courageous	○	○	○	○	○
Compassionate	○	○	○	○	○
Decisive	○	○	○	○	○
Emotional	○	○	○	○	○
Intelligent	○	○	○	○	○
Innovative	○	○	○	○	○
Organized	○	○	○	○	○
Sensitive	○	○	○	○	○

* 10. In your opinion, which one of the roles below would fit the agent you have seen in the picture above?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Personal Assistant	○	○	○	○	○
Health Counsellor	○	○	○	○	○
Exercise Coach	○	○	○	○	○
Financial Advisor	○	○	○	○	○
News Reporter	○	○	○	○	○
Storyteller	○	○	○	○	○

11. Please indicate why do you think you felt this way about this agent? This question is optional.

Embodied conversational agent survey questionnaire

* 12. In your opinion, which of these the following words defines the agent you see in the picture above?
The agent looks ...

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Affectionate	○	○	○	○	○
Ambitious	○	○	○	○	○
Aggressive	○	○	○	○	○
Creative	○	○	○	○	○
Courageous	○	○	○	○	○
Compassionate	○	○	○	○	○
Decisive	○	○	○	○	○
Emotional	○	○	○	○	○
Intelligent	○	○	○	○	○
Innovative	○	○	○	○	○
Organized	○	○	○	○	○
Sensitive	○	○	○	○	○

* 13. In your opinion, which one of the roles below would fit the agent you have seen in the picture above?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Personal Assistant	○	○	○	○	○
Health Counsellor	○	○	○	○	○
Exercise Coach	○	○	○	○	○
Financial Advisor	○	○	○	○	○
News Reporter	○	○	○	○	○
Storyteller	○	○	○	○	○

14. Please indicate why do you think you felt this way about this agent? This question is optional.

* 15. What do you think is the gender of this 3D character?
  Androgynous means having both female and male qualities.

Female   Androgynous   Male

* 16. What do you think is the gender of this 3D character?
Androgynous means having both female and male qualities.

Female	Androgynous	Male

* 17. What do you think is the gender of this 3D character?
  Androgynous means having both female and male qualities.

Female	Androgynous	Male