

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG



CHATBOTS' GENDER STEREOTYPES: INDIVIDUAL PERCEPTIONS OF THE GENDER OF CHATBOTS IN SOUTH AFRICA.

MASTER OF COMMERCE BY RESEARCH

FAITH NOKUTHULA NDALA, STUDENT NO.: 1419789 SUPERVISED BY DR. EMMA COLEMAN

ABSTRACT

Chatbots are getting closer to being integrated into our social environment and embodying our personalities, and how individuals react to information conveyed by chatbots is influenced by how they view chatbots. Research shows a small body of scholars studying individuals' views on chatbots and a minimal focus on gender. The views of gender affect how people interact with each other and potentially with chatbots. Although progressive steps towards gender equality have been taken in many countries over the last few decades, no country in the world has yet achieved gender equality. Considering the history of South Africa, gender imbalances date back to the apartheid years, when white men were primarily in leadership roles. Also, in the new age of democracy, women still suffer from the dominance of the patriarchy in the cultural, social, and political spheres. We may conclude that using chatbots with the unknown effects on how gender-based chatbots may reinforce gender inequalities in South Africa may increase inequality in the light of history.

The study's purpose was to explore the individual perceptions of the gender of chatbots in South Africa to raise awareness among chatbot designers of the implications that the gender of chatbots may have in society. To achieve the study's aim, the researcher adopted the Social Role Theory, Social Identity Theory, and Similarity-Attraction Paradigm to understand how gender roles are reinforced by people behaving according to them. In turn, these theories were used to address how people's perceptions of chatbots are influenced by its gender –by focussing on (1) how people respond emotionally in a controlled environment to chatbots of different genders performing the same role, (2) the influence of a chatbot's gender on the user's responses during the interaction, (3) the influence of a chatbot's gender on user's experiences, and (4) how people's prior-existing expectations of the role of the chatbot are affected by its gender.

Data were collected qualitatively through chatbot interaction and interviews. The researcher provided an experience for the participants by creating chatbots to assess the participants' experiences while communicating with chatbots of different genders and the same role and measuring how participants reacted emotionally to these chatbots in a regulated environment. The research design was based on a 2 x chatbot gender (male vs. female), 2 x participant gender (male vs. female), and 2 x stereotyped subject domain (auto mechanic and midwife) matrix. Following this, semi-structured interviews were conducted to obtain in-depth insights into the phenomenon while also aligning the results of the interviews with those gathered after the participants interacted with the chatbots. The participants interacted with the chatbots and were interviewed on the same day. The researcher was not present during the interactions with the bots, but interactions were

recorded for analysis. The researcher only joined the participants during the interviews, which were also recorded for analysis.

In South Africa, the results indicate that gender stereotypes in the offline world also exist in the online world and are more endorsed by males than females. People tend to associate feminine characteristics with female bots and masculine with male bots. These stereotypes are conditioned mainly by the societal norms developed to stipulate how people should act in societies and what work roles they should undertake. These attributes then make female bots more suited to providing health services and male bots to providing mechanic services. The female mechanic chatbot was more forgiven when it could not help with car diagnosis and was commonly perceived as incompetent. On the contrary, the perceptions about the male chatbot in the health sector were more about feelings around vulnerability.

While results revealed the same-gender preference as found in similar studies, this research also indicates cross-gender preferences such that some female participants preferred male bots, and male participants preferred female bots. The theory of social roles has emerged as more substantial than the other theories in defining and understanding gender stereotypes concerning chatbots. The findings also support Giddens's view that individual agency can reinforce and can change the social structure of gender stereotypes in this context. The use of Social Role, Social Identity Theories, and Similarity-Attraction Paradigm present more of an application of these theories to clarify chatbots' gender stereotypes where they have not been implemented in the past. The Structuration Theory brings together these ideas under the umbrella called structures to understand how gender roles are reinforced by people acting in accordance with them.

This study can pave the way for future research into how current social norms and stereotypes manifest in Artificial Intelligence interaction. In addition to methodological and theoretical contributions, the thesis can make practical insights into the decision-making process regarding chatbots' gender and Artificial Intelligence in general. This, in turn, may help drive gender equality through the use of attribute-based chatbots to facilitate improvements in the understanding of gender roles in South Africa.

ACKNOWLEDGMENT

This research would not have been possible without the exceptional support structure I had throughout completing it. I would like to thank my supervisor, Dr. Emma Coleman, for the ongoing encouragement, information sharing, and advice that has contributed a great deal from the time of the proposal to the submission of reports. I would like to thank Suzanne Sackstein and Nalukui Malambo, who have offered insights and experiences that have supported the research to test my instrument's validity.

A big thank you to Nicole Borges and Rodney Chetty for being incredible managers who have helped me from the beginning to the end of this research and given great advice, particularly on work-life balance and time management.

It is with great pleasure that I acknowledge the contributions of Bongani Motha, who gave up several evenings to read versions of this paper to validate ideas and grammar.

I am also grateful for the support of Christina Zachos, Akhona Mafenuka, Oduetse Nthetse, Nompumelelo Mphuthi, and Thuli Gwala.

I would also like to thank all the participants who availed themselves to be part of this study and those that helped during pilot testing. Also, to my classmates, you made the journey fun and entertaining. Thank you for your continued support.

Finally, to my siblings, Busisiwe, Thokozani, Mandla, and Thembi Ndala, thank you for your vows, trust in me, and assurance that I achieve my goal. Your efforts still leave me astounded.

PLAGIARISM DECLARATION

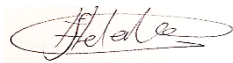
I, Faith Nokuthula Ndala (Student number: 1419789), am a student registered for MCom by research in the year 2018. I hereby declare the following:

I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.

I confirm that ALL the work submitted for assessment for the above course is my own unaided work except where I have explicitly indicated otherwise.

I have followed the required conventions in referencing the thoughts and ideas of others.

I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.



Signature: _____

Date: 19 April 2020

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGMENT	iii
PLAGIARISM DECLARATION	iv
TABLE OF CONTENTS	v
FIGURES AND TABLES	viii
1. INTRODUCTION	9
1.1. RESEARCH QUESTIONS	13
1.2. OBJECTIVES	13
2. LITERATURE REVIEW	15
2.1. CHATBOTS	15
2.1.1. DEFINITION	15
2.2. GENDER STEREOTYPES IN THE OFFLINE WORLD	16
2.2.1. GENDER STEREOTYPES	17
2.2.1.1. GENDERING AT WORK	18
2.3. GENDER STEREOTYPES AND SERVICE QUALITY: A CUSTOMER-TO-ATTENDANT ENCOUNTER. 21	
2.3.1. CUSTOMERS APPLY GENDER STEREOTYPES ON SERVICE DELIVERY	21
2.4. GENDER STEREOTYPES IN AFRICA	23
2.4.1. GENDER STEREOTYPES IN SOUTH AFRICA	24
2.5. CURRENT STUDIES ON CHATBOTS' GENDER STEREOTYPES	27
2.6. SUMMARY	29
3. THEORETICAL FRAMING	32
3.1. SOCIAL ROLE THEORY	32
3.1.1. ORIGINS OF GENDER DIVISION OF LABOUR	35
3.2. SOCIAL IDENTITY THEORY	36
3.2.1. THE EFFECT OF ATTRACTION: SIMILARITY-ATTRACTION PARADIGM	38
3.3. SUMMARY	39
4. METHOD	41
4.1. DATA COLLECTION STRATEGY	41
4.2. DATA COLLECTION INSTRUMENTS AND DESIGN	43
4.2.1. EXPERIENCE DESIGN	43
4.2.1.1. ENGAGEMENT FLOWS	47
4.2.1.2. CHATBOT FEATURES	48
4.2.1.3. TOOLS FOR CHATBOTS DEVELOPMENT	50

4.2.2.	INTERVIEW DESIGN	51
Table 3:	Interview questions.....	53
4.3.	SAMPLING AND SAMPLE SIZE	40
5.	DATA ANALYSIS PROCESS	42
5.1.	GENERATING THEMES.....	43
5.1.1.	CODING OF THEMES.....	43
5.2.	THEMES	44
6.	RESULTS ANALYSIS.....	46
6.1.	DESCRIPTIVE ANALYSIS.....	46
6.2.	ANALYSIS AND INTERPRETATION	50
	<i>THEME 1: PERCEPTION OF THE USE OF THE SYSTEM.....</i>	<i>50</i>
	<i>SUB-THEME 1: PRACTICE AND USE</i>	<i>50</i>
	<i>THEME 2: GENDER STEREOTYPES.....</i>	<i>52</i>
	<i>SUB-THEME 1: GENDER ROLE BELIEFS</i>	<i>52</i>
a.	INTERSECTIONALITY	53
i.	GENDER-BASED COMPETENCY	54
b.	RELIGION	57
c.	SOCIETY AND CULTURE	57
	<i>SUB-THEME 2: CONSCIOUS BIAS.....</i>	<i>60</i>
	<i>SUB-THEME 3: GENDER-BASED CHARACTERISTICS</i>	<i>61</i>
	<i>SUB-THEME 1: PERCEIVED DIFFERENCE.....</i>	<i>63</i>
	<i>SUB-THEME 2: SERVICE EVALUATION - BOT COMPETENCY</i>	<i>65</i>
	<i>SUB-THEME 3: PERCEPTION OF HUMANNESS</i>	<i>65</i>
	<i>SUB-THEME 4: VOICE</i>	<i>66</i>
	<i>THEME 4: EXPECTATIONS.....</i>	<i>67</i>
	<i>THEME 5: EMOTIONS</i>	<i>71</i>
	<i>SUB-THEME 1: COMFORTABLE.....</i>	<i>71</i>
	<i>SUB-THEME 2: IMPERSONAL.....</i>	<i>72</i>
	<i>SUB-THEME 3: FRUSTRATION.....</i>	<i>73</i>
	<i>THEME 6: RESPONSES</i>	<i>74</i>
	<i>SUB-THEME 1: RESPONSES THAT WERE THE SAME</i>	<i>74</i>
	<i>SUB-THEME 2: RESPONSES THAT WERE DIFFERENT</i>	<i>74</i>
	<i>THEME 6: SERVICE DELIVERY GENDER PREFERENCE.....</i>	<i>78</i>
	<i>SUB-THEME 1: HUMAN-TO-CHATBOT INTERACTION.....</i>	<i>78</i>
	<i>SUB-THEME 2: HUMAN-TO-HUMAN INTERACTION</i>	<i>80</i>
	<i>THEME 7: CHANGE DESIRED</i>	<i>83</i>

6.3.	GENERAL PERCEPTIONS.....	84
6.4.	SUMMARY OVERVIEW-WOMEN VS MEN VIEWS	86
6.5.	STEREOTYPE-FIRST VS. REVERSE STEREOTYPE-FIRST.....	88
7.	DISCUSSION OF THE FINDINGS.....	89
8.	CONCLUSION	96
9.	REFERENCE.....	107
	ANNEXURE 1: PARTICIPATION SUMMARY AND SCHEDULE.....	121
	ANNEXURE 2: TRANSCRIPT EXAMPLE.....	122
	ANNEXURE 3: CHATBOT INTERACTION EXPERIENCE EXAMPLE.....	125

FIGURES AND TABLES

Figure 1: “Operating mode of chatbots” (Zumstein & Hundertmark, 2017, p.98).....	16
Figure 2: The interaction of participants with both the male and female version of one of the roles.	45
Figure 3: Engagement flow between the participant and an auto mechanic chatbot.....	47
Figure 4: Engagement flow between the participant and a midwife chatbot.....	48
Figure 5: Basic architecture of the conversation between users and the chatbot.	49
Figure 6: Flow of conversation within DialogFlow.....	49
Figure 7: A chatbot deployed on Dialogflow Web.....	50
Figure 8: Themes after data collection.	45
Figure 9: Bots’ greeting message examples.....	51
Figure 10: P15 conversation example.....	57
Figure 11: P15 conversation example.....	76
Figure 12: P3 conversation example.....	77
Figure 13: Participants’ themes categorized by gender.	86
Figure 14: A proposed framework to explore the gender stereotypes of chatbots.....	95
 Table 1: <i>High-level Interaction Process</i>	44
Table 2: <i>Actual Interaction Process</i>	45
Table 3: <i>Interview questions</i>	53
Table 4: <i>Descriptive data</i>	49
Table 5: Methodological Justification Adapted from Klein & Myers (1999) Principles.....	104
Table 6: Participation summary and schedule	121

1. INTRODUCTION

The terms chatbot and conversational agents are sometimes used interchangeably by scholars to refer to the same thing, but they mean different things (AbuShawar & Atwell, 2015; Radziwill & Benton, 2017). Conversational agents are a form of a dialog system, such as Interactive Voice Response, that is often introduced for call centres (e.g., “Press 1 to speak to an agent”) that use a decision tree to communicate with a user. Conversational agents use cognitive tools such as Natural Language Processing to simulate human experiences and include agents such as chatbots (AbuShawar & Atwell, 2015; Radziwill & Benton, 2017). A chatbot receives the user’s natural language input as plain text or voice and then executes commands to engage to achieve the user’s intended target (Gustavsson, 2005). When the data arrives in the form of a voice, a speech recognition program is used to interpret the information (Mullennix et al., 2003). The newer developments of chatbots use machine learning to learn from and adapt to further the growing data.

Chatbots are not a new phenomenon. However, developments such as big data and machine learning have increased the use of chatbots. The early introduction of chatbots was through the development of ELIZA. The goal of developing ELIZA was to show that computers can understand the natural language and respond to the input of the user (Shum et al., 2018). The ultimate aim was to pass the Turing Test, which specifies whether a machine’s intellect is indistinguishable from humans (Radziwill & Benton, 2017). ELIZA used keyword matching to understand and react to user input, and with this limited intelligence, ELIZA could not keep the conversation going (Shum et al., 2018).

Another development of the chatbot was ALICE, which became famous not because of its conversational capabilities but Mark-up Language. Ding and Liu (2010) state that a “mark-up is regarded as a subtype of annotation and is defined as a formally structured annotation for a purpose, normally to allow some kind of manipulation of the information entity” (p.76). ALICE gave birth to what is known as Artificial Intelligence Mark-up Language (AIML), which was used to define a set of rules to match patterns from inputs. With this, ALICE was able to link users’ input with pre-defined topic categories. For example, an input “How do I pay through Electronic Funds Transfer (EFT)?” would be linked to a topic “EFT Transactions” (AbuShawar & Atwell, 2015).

In this study, the terms bots and chatbots are used interchangeably to refer to the same thing. The literature describes the bot as a computer program that performs automated repetitive tasks and classifies the chatbots as a type of bot (Radziwill & Benton, 2017). For this study, however, the word bot was used as a shortened version of the chatbot. In general, chatbots and robots are used

interchangeably to refer to the same thing. These words have been handled differently for this study. Although there are contradictory definitions of the term robots in the literature, the latter defines both robots and bots as programmable machines that can automatically execute actions, and the difference is that robots are with bodies and bots are programs (Appelbaum & Nehmer 2017). The definition adopted for this research is Alaieri and Vellino (2017), which describes robots as physical objects designed to perform a physical function. However, the difference between these words is that a chatbot can be created entirely from the software without any physical hardware (Alaieri & Vellino, 2017). It is also important to note that some of the participants referred to the chatbots as robots in their responses, and this may be because they may not know the difference between the two definitions.

South African consumers are making rapid use of digital media networks to avoid call centres and emails. As the number of mobile devices is growing in South Africa, Deloitte (n.d.) reports that most people frequently use their phones. This increase in use has made mobile phones the primary means of communication and marketing (Deloitte, n.d.). Smartphone adoption has been a key driver for chatbots (Brandtzaeg & Følstad, 2017), with today's consumers rapidly favoring digital platforms and a variety of instant messaging applications as their primary means of communication (Deloitte, n.d.).

Chatbots are moving closer to being incorporated into our social world and impersonating our identities. For example, in South Africa, ABSA has used chatbots on social media sites such as Facebook, Twitter, and WhatsApp to provide financial services to its clients (Moyo, 2017).

Anthropomorphism has been a successful strategy for integrating chatbots into our social environment and identities (Duffy, 2003; Zambaka et al., 2006). Anthropomorphism is the state in which human-like traits, both appearance and behavioural, are attributed to objects. One example is to attribute a chatbot with various emotions, such as being happy or sad in different social situations. This research focused on behavioural rather than a visual aspect of anthropomorphism (see section 4 for more details).

This research focuses on individual perceptions of gender. The words 'sex' and 'gender' are mostly used interchangeably (Torgrimson & Minson, 2005). However, literature has argued that they mean different things. "Sex is usually associated with biological differences and gender is associated with behaviour, social, cultural differences (such as a pink colour perceived as a colour for girls and blue for boy)" (Marcus, 2015, p.253). For this research, the words were treated differently. The word used is gender, as described by Marcus (2015), and only two genders, male and female, were considered for this research.

Little is known about individuals' perception of the gender of chatbots, and to expand existing research into a relatively new context, this research explored gender stereotypes by focusing on how individuals perceive the gender of chatbots. Studying this phenomenon is essential because "while positive steps towards gender equality have been taken in many countries over the last decades, no country in the world has yet achieved gender equality" (Ebert et al., 2014, p.359). Given the history of South Africa, gender imbalances date back to the apartheid years, when white men mainly occupied leadership positions. Even in the current democratic era, women still suffer from patriarchy's dominance in the economic, social, and political spheres (Booyesen & Nkomo, 2010; Ebert et al., 2014).

How individuals respond to information transmitted by chatbots is affected by how they interpret chatbots (Siegel et al., 2009). Research indicates a limited body of scholars studying individuals' views on chatbots, and minimal emphasis on gender; and views of gender affect how people communicate (Carpenter et al., 2009). It is then necessary to consider gender perceptions and their implications when designing chatbots that are to be employed in a social setting (Carpenter et al., 2009). Chatbots should be prevented from influencing individuals negatively, such as reinforcing gender stereotypes (Gockley et al., 2007). To practically avoid this outcome, it is crucial to understand humans' perceptions of chatbots' gender. Theoretically, "knowledge of how humans perceive and respond to robots can teach us much about human psychology, amongst other things" (Siegel et al., 2009, p.2563). We can conclude that employing chatbots with unknown effects of how the gender of chatbots may perpetuate gender stereotypes in South Africa may increase inequality given the history (Radziwill & Benton, 2017).

Therefore, this study aimed to explore the phenomenon of chatbots' gender stereotypes in South Africa using chatbots of different genders performing the same role with the same content to raise awareness of gender stereotypes in chatbots' design. Social Role Theory, Social Identity Theory, and Similarity-Attraction Paradigm were adopted to examine this phenomenon. These theories are used and proven in several gender studies and were adopted to see whether they hold validity for human-to-chatbot interactions. The data indicate that these theories hold validity in the online world; however, some elements arose from the data that the theories did not address.

The functions of chatbots selected for this research were auto mechanics and midwife agents. One male and one female chatbot provided auto mechanic services, and the others provided midwife services. The gender of the chatbots was represented by two attributes: voice (female vs. male voice) and names (typical female vs. male name). The data collection sessions were divided into two

phases. Phase I was the interaction session between the participants and the male and female chatbots of the same function, and phase II was the interview. The total number of participants was sixteen, eight males and eight females. The participants underwent both phases sequentially on the same day, based on the schedule (Appendix 1). Participants were chosen based on their availability, willingness to participate, and theoretically. The researcher chose an equal amount of female and male participants. Social media platforms, namely Facebook, WhatsApp, and LinkedIn, were the primary media used to invite participants. Data was collected qualitatively through chatbot interactions and interviews, which were digitally captured and analysed through thematic analysis.

1.1. RESEARCH QUESTIONS

RQ1. How are people's perceptions of chatbots influenced by its gender?

RQ.1.1. How do people respond emotionally in a controlled environment to chatbots of different genders performing the same role?

RQ.1.2. What is the influence of a chatbot's gender on the user's responses during the interaction?

RQ.1.3. What is the influence of a chatbot's gender on user's experiences?

RQ.1.4. How are people's prior-existing expectations of the role of the chatbot affected by its gender?

1.2. OBJECTIVES

The main objective is to explore how people's perceptions of chatbots are influenced by its gender.

The sub-objectives are to:

RO1. Determine how people respond emotionally in a controlled environment to chatbots of different genders performing the same role.

RO2. Identify the influence of a chatbot's gender on the users' responses during the interaction with a chatbot.

RO3. Identify the influence of a chatbot's gender on users' reported experiences with a chatbot.

RO4. Identify how people's prior-existing expectations of the role of the chatbot are affected by its gender

The results suggest that the gender stereotypes of chatbots exist and are conditioned mainly by the societal norms developed to stipulate how people should act in societies. People's assumptions of the gender of the bots in midwives and mechanic roles support the literature's statements. Such communal attributes like warmth and empathy are associated with females, and agentic attributes like being direct are associated with males. These attributes then make female bots more suited to providing health services and male bots to provide mechanic services. The female mechanic chatbot

was more forgiven when it could not help with car diagnosis and was commonly perceived as incompetent. On the contrary, the perceptions about the male chatbot in the health sector were more about security and weakness. Concerning chatbots, Social Role Theory and Giddens's view about agency came out stronger in explaining how gender roles are reinforced by people behaving according to them and how stereotypes change when people do not act according to them than the other theories.

The key obstacles were minimal literature and theories in Information System to investigate the phenomenon. The researcher has gained expertise from other fields, such as psychology and marketing. Also, the researcher designed the bots from scratch and did not use the available pre-built bots. For this study, it was essential to select stereotypic roles to fulfil the objective; however, pre-built bots were not suited to achieve this. Due to time constraints, the bots were not fed enough data to allow the participants to have more detailed conversations with the bots.

The next section summarizes the existing knowledge about what has been written about chatbots' gender stereotypes by also addressing the research questions' existing content.

2. LITERATURE REVIEW

This section is divided into four sub-sections. The first sub-section discusses chatbots as defined by the literature. The second sub-section looks at gender perceptions in the offline world. The section summarizes and gives an overview of gender stereotypes in human-to-human interaction. It highlights how people interact with one another and consciously or unconsciously apply gender stereotypes. Gender stereotypes and their implications are defined. The third sub-section summarizes the literature on gender stereotypes from a customer-to-service provider interaction perspective in the offline world. The section gives an overview of possible gender stereotypes in the offline world and which may apply in the online world. The fourth sub-section consolidates evidence of gender stereotypes in Africa. The final sub-section summarizes the literature on gender stereotypes in the online world, in human-to-computer interaction.

2.1. CHATBOTS

This section summarizes the definition of chatbots from the literature and discusses the technologies that enable a chatbot to understand the natural language. There is much debate about whether chatbots and virtual assistants are the same things or different. The last section consolidates evidence on how the two terms are treated, referencing the literature.

2.1.1. DEFINITION

Literature has classified chatbots as types of social robots developed to perform various roles that humans initially perform. A chatbot is a software program used to convey information using a Natural Language Processing Engine (NLP) to understand the natural language or use keywords from the input and retrieve the information that matches the keywords or pattern (Radziwill & Benton, 2017). The chatbot comprehends the request made by users expressed in natural language, and it can react and respond like humans. It is as if there is a person behind the computer; however, it is the programmed chatbot with human-like traits portraying illusion (Gustavsson, 2005). The information conveyed by the chatbots can be text or voice. Text-to-speech (TTS) conveys the text to spoken voice response (Mullennix et al., 2003).

These chatbots are designed to communicate for a specific purpose and not for the entire range of human conversations. The conversation skills are improved continuously through machine learning capability, which gives chatbots the ability to learn through data. (Haung et al., 2007; Shawar & Atwell, 2007).

Figure 1 shows a high level of a chatbot's operational model from when a user sends a request to when the chatbot responds with an answer. The process starts with the user sending a request to a chatbot using messaging platforms such as messenger or speech input, such as Amazon echo, as illustrated in step 2. The Natural Language Parser then records and interprets the user's natural language input into a language the conversation engine understands. The conversation engine analyses the request and sends it to the back-end where different commands run to pull back a response from information warehouse such as databases.

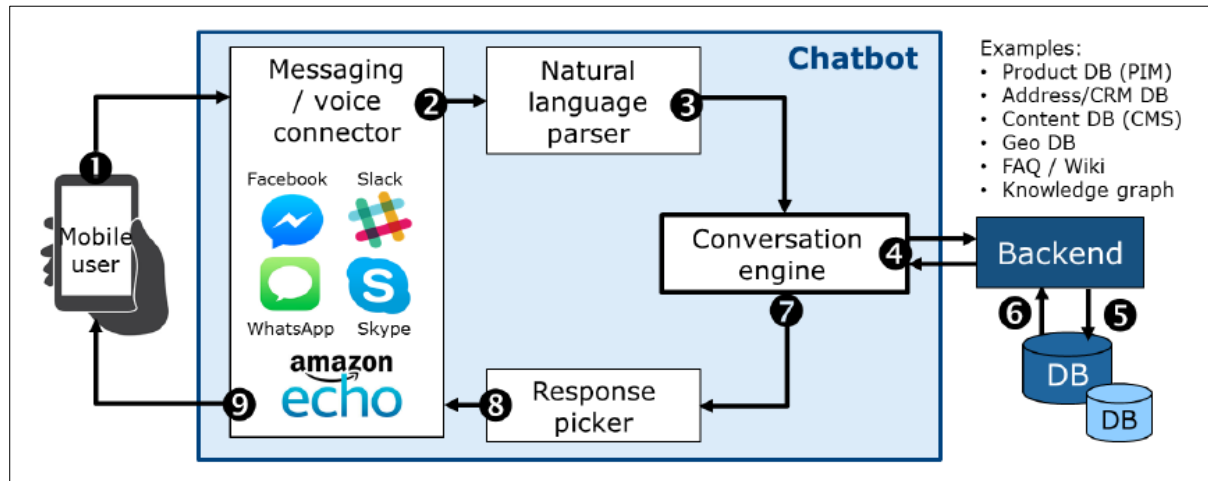


Figure 1: "Operating mode of chatbots" (Zumstein & Hundertmark, 2017, p.98).

Some chatbots are presented as females, males or are kept gender-neutral (Marcus, 2015). Historically, secretaries and administrative assistants are women; therefore, in most cases, the default gender prescribed to a chatbot whose tasks are to perform the day-to-day tasks of secretaries or administrative is female. When developers design a chatbot to automate a task, they tend to have an idea in their minds of the humans who were performing the tasks before then. The default gender becomes that of those humans. Chatbots with human-like traits can perpetuate stereotypes (Dinat, 2017). For example, Alexa is a female, but why? What messages are being reinforced by Alexa? There seems to be no literature addressing how the gender of chatbots' is selected. This is discussed in more detail in section 2.2. It becomes easy to identify the gender of the chatbot given the characteristics such as the bot's voice (Chaves & Gerosa, 2019). De Angeli et al. (2001), as stated by Chaves and Gerosa (2019), claim that "a chatbot's body can be created through the narrative without any visual help and talking to a machine affords it a new identity" (p.25).

2.2. GENDER STEREOTYPES IN THE OFFLINE WORLD

Research in Information Systems shows a small body of scholars investigating individuals' perceptions towards robots and very little focus on chatbots' gender. Literature was drawn from

marketing, psychology, and sociology studies and used to understand the phenomenon and outlines the implications of chatbots' gender stereotypes. This section summarizes gender stereotypes in the offline world. For this research, the offline world refers to human-to-human interaction, while the online world refers to human-to-chatbot interaction. Gender stereotypes can occur both at home and at work (Heilman, 2001). For this research, this segment aims to provide awareness of gender differences at work since the chatbots' tasks are performed at work rather than at home.

2.2.1. GENDER STEREOTYPES

"Stereotypes are sets of beliefs about other people reflecting shared generalisations about members of a social group. Stereotypes act as 'default settings' in social perception, providing a 'best guess' about an unknown group member. When people first meet, stereotypes are heavily applied but eventually give way as more individuating information is obtained" (De Angeli & Brahnam, 2006, para.2). Many people use their gender to define traits such as who they are and what they believe. "Gender, of course, is a form of human variation that is highly susceptible to cultural generalization as a primary category for framing social relations" (Ridgeway, 2009, p.148). Thus, the inherent differences between women and men are a primary cultural category used to categorize people. For example, in a social interaction setting, an individual almost instantly gender categorizes others to identify similarities or differences to relate (Ridgeway, 2009). This act unconsciously triggers stereotypes and influences how people behave afterward.

Literature stipulates that the act of categorizing people on gender is not only evident in a human-to-human interaction setting but also on the Internet and imaginatively (Chattaraman et al., 2011; Siegel et al., 2009). For example, in recruitment, one may categorize a person after examining the resume, and others create an imaginary person they will want to hire. These categories create "shared cultural beliefs that stipulate how some people in one category behave compared with the opposite" (Ridgeway, 2009, p.148). Even those who do not hold those beliefs expect to be judged because most people hold these beliefs. Therefore, they consider these beliefs in their behaviour (Ridgeway, 2009). Kulms et al. (2011) state that women are likely to be negatively judged when they act against the stereotypes the societies hold about them compared to men. For example, with non-verbal behaviour, women are perceived to show more intimacy than men. If a woman does not smile, she evokes negative evaluation. However, men are not expected to fulfill comparable behaviour expectations because, in their case, these cues are not existing. "Thus, men are not perceived to reveal a negative emotional state if they do not smile, as non-smiling men are seen as the norm, whereas women are perceived to deviate from the norm when they are not smiling" (Kulms et al., 2011, p.82).

Using gender to differentiate between people leads to gender stereotypes. "Gender stereotypes are our beliefs about how most people view the typical man or woman" (Ridgeway, 2009, p.148). Literature states that gender stereotypes stem from a long time ago, and even in changing societies, gender stereotypes remain firm. This evidence is apparent in the study done by Ridgeway (2009) to consolidate the evidence of the ongoing inequality of women in the workplace in the United States. Ridgeway (2009) argues that gender issues are not solely the past or cultural lag results, but they are re-created daily (Blair-Loy & Herron, 2012). This notion results from continuous, shared social expectations. According to Social Role Theory, which is discussed further in section 3, gender stereotypes describe how women and men behave and prescribe how they should behave (De Angeli & Brahnam, 2006; Heilman, 2001; Knobloch-Westerwick & Hoplamazian, 2012). Prescriptive stereotypes rise from social scripts that are learned by individuals in a society that, in turn, enforces the stereotypes by regulating communication between members of the same or different gender (De Angeli & Brahnam, 2006).

2.2.1.1. GENDERING AT WORK

Occupational role stereotyping, which is an act of associating jobs with gender, is a global issue that most organizations are still battling to end (Baker, 2014). Some organizations are continually seeking ways to challenge internal inequality and promote diversity. For example, in Information Technology (IT), there are programmes such as FemBioBiz developed to impact women in IT and elevate them to top managerial positions (Letsebe, 2018).

The literature shows an increased number of women in the workplace over the years to achieve gender equality. Kerevel and Atkeson (2015) state that gender equality improvements can either change or reinforce gender stereotypes. For example, Kerevel and Atkeson (2015) have found that the presence of a female mayor in Mexico has dramatically modified men's perceptions that women cannot be competent leaders and that they have agreed to be led by a female mayor. Eagly and Wood (2012) state that an increased presence of women in the workforce occupying roles perceived as 'male roles' changes gender stereotypes. Pratto et al. (2006) argued that shifts in gender inequality do not change gender stereotypes but instead bring about shifts in power so that women in leadership positions have the same power as men in the same position. This is evidence that both females and males face gender stereotypes (Stamarski et al., 2015). Mostly men occupy leadership positions, and most nurses are females. This section highlights that gender role is not one-sided; the victims are not only women but also men.

The focus of this section is on occupational role stereotypes in leadership positions and the nursing field. These roles are commonly stereotypical such that nursing is perceived as the females' sector and perceptions that most men than women occupy leadership positions. For this research, this section's primary goal is to highlight that both genders, males and females, face gender stereotypes.

2.2.1.1.1. MEN OUTWEIGH WOMEN IN LEADERSHIP POSITIONS

Baker (2014) conducted a study to provide the current trends of women in leadership positions. The study's purpose was to provide a literature review that stipulated that even though women can be top performers, they are still disadvantaged in obtaining leadership positions compared to their male peers. Women are more disadvantaged than men for career advancement opportunities (Ebert et al., 2014; Kiser, 2015). There is an ongoing debate as to why the imbalance persists. Even though there is evidence of women's presence in middle managerial positions, women are still bottlenecked for promotion, whereas men's opportunities to be promoted are clear and easily attainable (Kiser, 2015). The culture of gender inequality appears to be promoted by men in the executive positions who seem less concerned about advancing women to top managerial positions (Kiser, 2015; Stamarski et al., 2015). Baker (2014) postulates that managers fail to promote women into senior positions because they believe that women do not possess enough agentic characteristics such as strength and technical abilities.

The lack of fit model developed by Heilman in 1983 and 1995 explains this notion clearly. The imbalance caused by the perception of women attributes against required attributes associated with male-type roles creates a perception that women lack the fit in work settings. The perceptions of how a person will perform in a job drive personnel decision. If a person's attributes match the job requirements, the fit is good, and success is expected; failure is expected if the fit is poor. It is clear from the literature that the attributes for managerial roles, such as decision-makers, assertive do not match those believed to categorize women, such as empathy and caregiving. Therefore, the lack of fit of these stereotypical attributes is likely to drive expectations of failure. These expectations are portraying women as incompetent to succeed in male-dominated roles (Heilman, 2001). Women are expected to work hard, twice as men, to prove that they are worthy of the managerial positions (Horwitz et al., 2018; Selzer & Robles, 2019). In Kiser (2015) research, his findings indicate that men feel they have more right to a job than women when jobs are scarce. Also, he postulates strong feelings that children suffer when a mother works for pay to provide for them and that in this regard, men make better leaders.

The social dominance theory gives a different perspective on why people dominating are less concerned with advancing others. The theory postulates that people are divided into hierarchies, and those dominating do what they can to stay dominant. It explains why men in executive positions are less concerned about advancing women. Theoretically, if women were to dominate in executive positions, they would do the same (Pratto et al., 2006).

2.2.1.1.2. MALE NURSES FACE GENDER-BASED OCCUPATIONAL BARRIERS

Mostly, nurses are females, which stems from professors and textbooks mispresenting gender diversity in the nursing field by referring to nurses solely with feminine pronouns (Escobar & Heilemann, 2019). Media matters also portray only women as nurses (Escobar & Heilemann, 2019). This stereotype has led societies to exclude men in the nursing field socially. It seems that people have ignored the history of the role men played in the nursing field (Clow et al., 2014). It is evident in the literature that historically, men occupied the nursing field more than women mainly because of the military. Only men occupied the first nursing school in India. Back then, in the 1990s, men were considered competent for the role. They had been involved in remarkable historical incidents where they helped the injured and the sick during the Crusade in the 11th century (Ross, 2017).

This perception of nurses being only females has led people to be less accepting of male nurses (Clow et al., 2014). For example, Taiwan male nurses have reported that when interacting with their patients, they have to continually convince them that they are real nurses (Yang et al., 2004). Males pursuing nursing as a career may be perceived as pursuing a career that is incompatible with their gender, which enforces gender stereotypes, and this is evident in the study conducted by Harding (2007). The males in New Zealand have reported that males' public perception in nursing has labeled them as gays even though they are heterosexuals (Harding, 2007). This notion prevents male nurses from entering the nursing field and building relationships with their patients (Evans, 2002).

2.3. GENDER STEREOTYPES AND SERVICE QUALITY: A CUSTOMER-TO-ATTENDANT ENCOUNTER.

The use of chatbots is to render services to customers over the Internet. Since the literature lacks studies on gender stereotypes from a customer to service provider online, this section focuses on consolidating a perspective from the offline world between a customer-to-attendant encounter (an encounter between the customer and the service provider). For this research, the service providers are the chatbots since they offered services to the participants.

2.3.1. CUSTOMERS APPLY GENDER STEREOTYPES ON SERVICE DELIVERY

Organizations are continually seeking ways to create a competitive edge over their competitors. Service quality has been vital in creating a competitive advantage for organizations, and it is sought after by many organizations (Deshwal, 2015). The quality of the service can be determined by how well the service delivered meets the needs of the customer (Sayareh et al., 2016). This section summarizes the literature on gender role stereotypes applied by customers when measuring service quality.

Studies have reported that customers' service evaluation is affected by the service provider's gender (Luoh & Tsaur, 2007; Snipes et al., 2006; Wang, 2016). Customers apply gender stereotypes that one gender is better at delivering a given service than the other in different service environments. Pinar et al. (2017) study's key objective was to examine the gender effect to decide whether consumers receive better service from the same-gender service providers, as suggested by Similarity-Attraction Paradigm, or they receive better service from opposite-gender service providers, as suggested by flirting theory across four different cultures. The findings report that female respondents have a significantly higher perception than male respondents that males are better off than females in providing a better quality of low-status services. In comparison, male respondents have a significantly higher perception than female respondents that males are better off than females in providing a better quality of high-status services. Low-status services in this context refer to low-skill services such as call centre agents, and high-status services refer to high-skill services such as technicians.

On the contrary, Mattila's (2000) findings indicate that it is instead culture bounds than gender differences that influence customers' service evaluation. Most customers' perceptions are based on the service provider's physical appearance, their pre-established expectations, and where the interaction took place (Luoh & Tsaur, 2007). When the information about a service provider's identity is observable, individuals use the observable cues to differentiate the service providers and

form impressions. These differences can be observable social features such as race, age, gender (Bartlett & Gulati, 2016).

The gender-based occupational exclusion by customers leads to stereotypes that one gender is better at delivering a given service than the other (Luoh & Tsaur, 2007). In most sectors, service employees have increased their emotional investment in customer service due to the increasing demand for personalized services and quality. For example, the front-office services where employees emotionally bond with customers are occupied by females. This was found to be true in the airline and with Mc Donald's window workers. Leider's (1993) study, as stated by Gustavsson (2005), stipulated that male workers at Mc Donald's found women to be more competent in customer relations than them. These assumptions about women's competency also prevail in the airline industry, which leads to stereotypes as in pilots are men and flight attendants are women.

Literature states that there are at least two kinds of stereotypes in a service environment setting; they are server gender stereotypes and in-group stereotypes. The server gender stereotypes are expectations as to what the gender of the service provider should or will be in a particular environment (De Angeli & Brahnham, 2006; Mulvey & Killen, 2015). The in-group stereotype in the customer-attendant encounter refers to the expectations that individuals will receive better service delivery from the service providers of the same gender (Luoh & Tsaur, 2007). Pinar et al. (2014) argued that the perception of customers' service quality could be "affected more by occupational stereotypes than the gender of the service provider" (p.151). For example, a customer's perception of service quality could be affected by males in nursing rather than the nurse being male or female. A study was conducted on multi-dimensional industries to determine the "effects of gender similarity/dissimilarity, gender stereotyping, and culture on perceived service quality" (Pinar et al., p.149). The study found that gender differences between the customers and the service providers did not significantly affect the service's perceived quality. However, occupational stereotyping and the differences in culture had a significant impact.

2.4. GENDER STEREOTYPES IN AFRICA

This section summarizes the evidence of gender stereotypes in other African countries other than South Africa. This illustrates that gender stereotypes are a world's issue, not only evident in South Africa.

Makama (2003) postulates that women in Nigeria have been side-lined to trading and selling goods on the streets or markets. Makama's (2003) study perceives Nigeria as a highly patriarchal country, which is a "major feature of a traditional society" (p.116). Like their counterparts in other countries, women in Nigeria face discrimination in opportunities to advance themselves. The domestic burden on their shoulders has led to their inputs being perceived as invaluable. This result is due to the reproduction burden in which women are expected to be active in "roles of a childbearing, child-raising, and homemaking" (p.122); one of the reasons why Nigeria has a high rate of childbirth (Makama, 2003). However, there has been an increase in women's participation in politics (Makama, 2003). This increase includes women being able to vote and political offices held by women, even though the number of men in politics still outweighs women in Nigeria.

Okafor et al. (2011) conducted an empirical analysis of barriers to women's leadership and managerial aspirations in Lagos, Nigeria. The study stipulated that women have the attributes to be in managerial positions; however, they are affected by family issues, individual and organizational factors (such as gender role stereotypes). A claim is made that a significant number of women are frustrated and feel trapped by the challenges of balancing work and their lives outside work. A lot of activities women have to perform are overwhelming and can create work-life conflicts. For Nigeria to combat this slow development, sustainable "focus should be on the education of the girl child and capacity building, to achieve economic empowerment of Nigerian women" (p.6724).

In Zimbabwe, changes in gender stereotypes were made possible by women organizations that were initially developed in 1980 to improve the welfare of women (Kurebwa, 2014; Chabaya et al., 2009). In those years, the organizations were not effective because they were put in the urban areas and isolated women in the rural areas where most gender issues arose (Chabaya et al., 2009). In 1990 emphasis on these organizations shifted from welfare to developing women, and women in rural areas were no longer isolated. These organizations then became effective, driving women empowerment (Kurebwa, 2014; Chabaya et al., 2009).

Kurebwa (2014) stipulated that women in rural areas in Zimbabwe are challenged by two significant factors that prevent them from accessing and participating in local governance. These factors are

under-representation and low participation compared to women in urban areas and men in general. The two factors result from barriers such as gender stereotypes, discrimination, lack of education, and culturally described domestic roles.

The study by Chabaya et al. (2009) is in agreement with Okafor et al. (2009) that some women do have attributes to be in leadership positions potentially but are affected by individual gender-imposed stereotypes. The study found that Zimbabwe's female teachers have adequate qualifications and attributes, which qualifies them for school headship in primary schools. However, most of these women never attempted to apply for these positions because of gender role stereotype cues that these women potentially inherited from their societies, which manifested in forms such as low self-esteem and lack of confidence. These women showed a preference for family responsibilities over their career advancement. If they were to choose career over family responsibilities, they would be overwriting the importance of being a mother or a wife first (Chabaya et al., 2009).

2.4.1. GENDER STEREOTYPES IN SOUTH AFRICA

This section summarizes the evidence of gender stereotypes in South Africa by focusing on leadership and nursing roles. This illustrates that gender stereotypes in leadership and the nursing field exist in South Africa, even in this era of democracy.

South Africa is one of the countries with many diverse cultures and religions. South Africa borders include the Zulu, Xhosa, Pedi, Tswana, Ndebele, Khoisan, Hindu, and Muslim people, to name but a few. All these people are united in calling South Africa home, and thus their lives all lead to being part of the country's history, identity, and culture. This led South Africa to be known as the Rainbow Nation. Even though South Africa comprises different people from different backgrounds, regarding gender roles, the belief is similar across cultures throughout history (Thoreson, 2008; Sidanius et al., 2019).

Looking back on South Africa's history of gender roles, women have traditionally been assigned to agriculture and domestic work and child care in rural communities. Men tended livestock, worked hard farm labor, and ran local politics (Hutson, 2007). With the African peasantry's dispossession, many men became migrant workers in remote employment centers, leaving women to manage rural households (Budlender & Lund, 2011). As time went by, women also became migrant workers with limited opportunities because of the apartheid government's restrictions. They were mostly performing domestic and casual work in farms owned by white people (Hutson, 2007). While women

and men were employed in urban areas outside the home, women were still responsible for domestic chores and child care. If women working outside the home in urban areas had children, these household duties generally fell under the responsibility of older female children, who had to manage those responsibilities and education. These gender roles in the history of South Africa, as in some countries, have socialized men as heads of households who often manage institutional capital (Hall & Posel, 2019).

The difference in the historical racial hierarchies under the apartheid system in South Africa has led to differences in gender and race experiences in the workplace. It is evident in South Africa that even though white women acquired 'white privileges' which presented them with more favourable opportunities than black women, they are also faced with barriers to male-dominated roles (Booyesen & Nkomo, 2010). On the other hand, African black women were domestic workers, and white men occupied the highest managerial positions. Although apartheid has ended in South Africa, and employment legislations are in place, women still suffer from the dominance of patriarchy in the economic, social, and political spheres. The hierarchies that existed in the apartheid years are still dominating in today's age. Whites still occupy the top managerial positions than black people (Booyesen & Nkomo, 2010).

It is only post-apartheid that the government of South Africa has begun to drive and affect change in gender roles. The new democratic constitution is based on global humanitarian values and promotes equality between men and women and other human rights (Hassim, 2003). For example, the Commission on Gender Equality is one of the institutions that promote constitutional democracy in South Africa. The Commission has been set up to promote gender equality, advise and make recommendations to the government on any legislation affecting gender equality and women's status. Although the government in South Africa is pushing these reforms, gender inequality continues to be prominent. Women still suffer from the supremacy of the patriarchy in the cultural, social, and political spheres (Van Der Heever et al., 2019).

Remaite (2013) conducted a study concerning women's perceptions and their experiences in leadership roles in South Africa, taking into account their gender. The study reports that women still face challenges in leadership roles. They are often disadvantaged over men for appointments or promotions to senior roles and have been unfairly regarded because of the perception that women are not suitable for these roles. This has prevented some of them from applying or putting themselves up for promotion in those positions. Instead, they prefer to stay in positions that they feel secure in and believe will allow them to work and maintain a balance of life, such as being

entrepreneurs. Remaite (2013) also notes that it is only in entrepreneurship that some women feel that they would not mind taking a higher role if they were in their own companies.

One would expect the Remaite (2013) findings to be different from more recent studies. However, there seems to be a prolonged increase in gender inequality in the workplace and evidence that even as women have successfully attained leadership positions, they are nevertheless judged negatively and discriminated against for career advancements. Alahdal (2018) states that there are gender gaps between women and men in senior positions at universities in the Gauteng province of South Africa and that institutions are still seeking to address these. More women than men reported being discriminated against in their positions, and these discriminations were not institutional obstacles but secretly applied. The greatest obstacle identified was the lack of support from family members and social expectations for women to take care of children, which hinders their career advancement. This brings significant difficulties in juggling work and life, as stipulated by Remaite (2013).

The majority of the studies conducted in South Africa concentrate on gender stereotyping of women, which is driven by history so that, considering the history of South Africa, women are more disadvantaged than men. This should not dispute that men are still discriminated against in some areas, such as nursing. Men are a minority in the nursing workforce in South Africa. In a study conducted by Ndou and Moloko-Phiri (2018), four-year Diploma male students reported experiencing discrimination in clinical settings. They were mostly assigned non-infirmar tasks, which made them feel like their learning has not been prioritized. However, their female counterparts were viewed favorably and were given more chances to practice nursing while carrying out tasks regarded as manly (Ndou & Moloko-Phiri, 2018), such as lifting heavy equipment and patients. In practice, men who are already working as nurses in South Africa have reported discrimination because nursing is viewed as a female occupation (Ndou & Moloko-Phiri, 2018). Kalemba (2020) states that the gender and race hierarchies of the apartheid period in South Africa are a central factor in the organisation and feminization of nursing in South Africa. Men's career preferences in female-dominated occupations result in them being branded as gays because the roles are considered only suitable for females.

In summary, the history of South Africa has affected differences in gender roles stereotypes. South Africa's colonial history influences how roles are perceived (Budlender & Lund, 2011). Evidence of gender stereotypes in leadership positions and nursing in South Africa indicates that South African history has a long-standing hierarchical gender structure in South African societies. These hierarchies

have shaped communities on how gender roles are subordinated and oppressed in the workplace in unique ways to the South African social context. It is important to remember that both men and women in South Africa face gender roles in various ways, and if these gender stereotypes are carried to the online world, it will be challenging to achieve gender equality in South Africa as this will reinforce gender stereotypes.

2.5. CURRENT STUDIES ON CHATBOTS' GENDER STEREOTYPES

This section consolidates evidence of chatbot gender stereotypes from the current studies on this phenomenon.

Chatbots are not a new technology; they have been developed in the past but have gained attention in the Information Systems research body (Shum et al., 2018). However, chatbots' gender stereotype is a new phenomenon about which little literature has been written. There is evidence in the literature that chatbots' gender stereotypes exist.

For example, Baxter et al. (2018) studied the impact of chatbot gender on users' stereotypical perceptions and satisfaction. Six chatbots (2 x male, 2 x female, and 2 x non-gendered) were developed to test this phenomenon in two occupational roles (banking and mechanics). Their finding indicates that chatbot gender influences "users' overall satisfaction and gender-stereotypical perception" (para.1). This theory suggests that users are more likely to apply gender stereotypes to chatbots that work within a gender-stereotypical subject domain and those that do not adhere to existing gender norms in societies. The non-gendered chatbot received positive feedback compared to the gendered bots. There were no significant discoveries that indicated gender stereotypes in a gender-neutral subject domain, such as banking. However, the participants applied stereotypes towards the female chatbot that was offering mechanic services.

Literature states that the human-like characteristics given to chatbots influence individual perceptions. As the chatbots' interface becomes more natural than before, the tendency to attribute chatbots as having human-like attributes increases (Holtgraves et al., 2007). Gustavsson (2005) states that the companies in his study have denied or been ignorant of the consequences of the chatbots' gender; however, they have indirectly expressed favoring female chatbots. For example, Humany, a Swedish software development company, has employed Doris and Aurora as site assistants, and the argument is that they could have chosen non-stereotypical names rather than female-stereotypical names Doris and Aurora.

How the chatbot looks influences the users' assumptions about traits such as gender (Baylor, 2009)? Literature indicates that chatbots from this stance can positively change the behaviour of individuals. For example, the study that was done by Baylor (2009) has reported a positive influence on students' behaviour toward engineering. The students have reported positively from engaging with a female chatbot offering engineering services. Baylor (2009) stipulated that this might be because the interaction might have challenged the students' existing beliefs about engineering.

The arguments about whether female or male users prefer female or male chatbots are conflicting in the literature. The studies stipulate different findings under a similar purpose of observing gender role impact on users' gender preference. The researcher argues that these studies are in different contexts, environments, populations, and there is less harmony in the conclusions. Most of the studies do not solely focus on gender; they have incorporated other virtual characteristics such as animation. For example, the study that was done by Nass and Moon (2000) postulates that both female and male users found female-voiced chatbots less friendly than the male-voice chatbots even though the content embodied in these chatbots was the same. This contradicts the offline world of females being more emphatic than men.

In contrast, Siegel et al. (2009) studied Persuasive Robotics and focused on the influence of robot gender on human behaviour. The study was about humanoid robots with a body shape designed to imitate a human body and not conversational bots. They adopted three measures to study people's perceptions of robots: trust, credibility, and engagement (Siegel et al., 2009). The study reports a cross-gender preference across these three measures that males preferred a female robot, and females preferred a male bot. However, "the donation measure deviates slightly from the cross-gender pattern seen in the subjective measures; men donate significantly more often to the female robot, but women show no preference. Men's donation behaviour remains consistent regardless of whether or not other museum visitors accompanied them. On the other hand, women donate more often to the female robot when accompanied, but reverse their preference to slightly favour the male robot when alone" (p.2567).

This result did come as a surprise as some "literature in social psychology would tend to suggest a same-gender preference rather than a cross-gender preference. This stems from a general tendency for people to be more easily persuaded by similar or in-group members. This tendency was found to be true in similar work with virtual humans in immersive virtual environments" (Siegel et al., 2009, p.2567). However, this was measured on one behavioural component of gender: a voice; other components beyond the voice changes might give different results. Individuals apply gender

stereotypes to chatbots and respond differently depending on the gender and the chatbot's role. Users apply social scripts of gender to computers. Mou et al. (2019) state that participants rated male voiced chatbots as "more friendly, convincing, competent, and knowledgeable on mechanical conversations, whereas the female-voiced chatbots were viewed knowledgeable on relations but yet less competent" (p.3).

In the literature collected by Gustavsson (2005), a female image was used to personalize the chatbots rather than a male image. Gustavsson (2005) stipulated that the reasons may be from the rooted stereotyped image of females as being pleasant or "it might be easier to design an idealized and acceptable image of a female than of a male because there are more 'cyberbabes' than 'cyberhunks' on the Internet to guide and inspire designers in their work" (p.415). This is evidence that gender stereotypes exist in the online world.

Gendering chatbots may perpetuate negative stereotypes about that gender. For example, Tay, the chatbot that Microsoft released on the 23rd of March 2016 on Twitter, has painted females negatively on Twitter (Abrahams, 2018). Microsoft shut down the bot after sixteen hours it has been up to because of the offensive discrimination posts it posted on Twitter. This tragedy raises concerns about what happened when we anthropomorphized technology. Professor Gina Neff stated that this tragedy could have been avoided if women have been involved in the concept realization of Tay the bot. She added that women could have given Microsoft team a better visual of how it is like to be a woman on Twitter, primarily as she has referred that Twitter is not a safe platform for women to raise their opinion and voice (Abrahams, 2018).

2.6. SUMMARY

The purpose of this review was to holistically view existing knowledge about gender stereotypes in the offline world and also to look at this phenomenon from an African and South African context. There is little research on chatbots' gender stereotypes; the review consolidated this phenomenon from existing research both from an online interaction between human-to-chatbots and offline interaction between human-to-physical robots. Also, exploring potential parallels with human-to-human interaction.

It is evident from the literature that gender stereotypes are widely articulated in human-to-human interactions. However, there is less evidence that these behaviors are taken into account in the online environment, particularly with chatbots. With these chatbots becoming more integrated into our lives and personalities, they are likely to perpetuate gender role stereotypes when gender differences are not removed in their design and development phases. Given that there is little

research on gender stereotypes of chatbots, this study needed to be carried out because gender expectations and their implications need to be considered when designing chatbots to be used in a social setting, particularly in South Africa.

The literature indicates that even after the apartheid regime in South Africa adopted a constitution to address gender discrimination, it is still struggling to achieve gender equality. The history of South Africa is one of the factors that led to long-standing perceptions of gender roles. If chatbots are employed in South Africa without understanding how chatbots gender roles can affect South Africa, gender stereotypes may be reinforced. This, in essence, would make achieving equality between men and women more difficult.

The literature gaps regarding chatbots' gender stereotypes have resulted in this paper's primary research question. Concerning the main research question, 'how are people's perceptions of chatbots influenced by its gender?' It is clear from the literature that in the offline world, communal attributes such as caring, empathy, nurturing, kind are associated with females, and agentic attributes such as assertiveness, strength, directivity are associated with males. We expected participants to assign communal attributes to the female bots and agentic attributes to male bots from these associations. Also, practical roles are presented as male-dominated, and roles that encompass the care of individuals such as nursing are presented as female-dominated. We expected the participants to associate the midwife's role with the female bot and the mechanic role with the male bot.

The sub-research questions were:

RQ.1.1. How do people respond emotionally in a controlled environment to chatbots of different genders performing the same role?

RQ.1.2. What is the influence of a chatbot's gender on the user's responses during the interaction?

RQ.1.3. What is the influence of a chatbot's gender on user's experiences?

RQ.1.4. How are people's prior-existing expectations of the role of the chatbot affected by its gender?

The review does not present a unified perspective about how peoples' perceptions of chatbots are influenced by its gender concerning emotions, influence on user responses, experience, and prior-existing gender role expectations. The literature revealed that individuals' interaction with the bots

might or not be affected by the bot's gender. We can expect that the gender of the bots may or not influence participants' responses and how the participants respond emotionally in a controlled environment to chatbots of different genders performing the same role. There is inconsistency in the reports of individuals' perceptions of the gender of the bots in reversed-stereotypical domains from the online world. Some reported that in practical roles, female bots were preferred (Nass & Moon, 2000), which contradicts gender role beliefs as stipulated by some parts of the literature, whereas some findings deemed them as incompetent, and the male bots were not deemed as incompetent (Baxter et al., 2018). Given the possibility that most participants' prior-knowledge about gender roles matches the claims in the literature, we expected that some of the participants would find it normal for the bots to operate in the unexpected role, and some would find it unusual.

3. THEORETICAL FRAMING

This section summarizes the theories that were adopted to carry out this research. These theories have created a roadmap that helped investigate the phenomenon to address the research questions. Owing to the absence of Information Systems theories to investigate the phenomenon, the researcher adopted theories from sociology. For this research, Social Role Theory for gender differences and similarities by Eagly and Wood, the Social Identity Theory by Tajfel and Turner, and Similarity-Attraction Paradigm by Byrne and Rhamey were adopted to explain how individuals apply gender stereotypes-how gender roles are reinforced by people behaving in accordance with them. These theories have been used in various gender studies in sociology to investigate stereotyping in human-to-human interaction. They are supported in the offline world and were applied to the online world of human-to-chatbot interaction.

3.1. SOCIAL ROLE THEORY

The theory of social roles was initially formulated in the 1980s to explain gender differences and social behaviour similarities. The Social Role Theory (SRT) suggests that the variations in men's and women's actions are caused by the contrast between men and women's distribution into social roles (Harrison & Lynch, 2005). Such functions are symbolic of how men and women behave and are prescriptive of how they should behave. Women are classified into feminine and men into masculine roles. The feminine role is followed by social traits concerned with the individual's well-being, such as empathy, whereas the masculine role is followed by agentic traits with bold and commanding tendencies (Frawley & Harrison, 2016). As a result, gender role roles are reinforced in both men and women so that men appear to be agentic, and women tend to be cooperative.

The theory of gender roles persists because of both males' and females' observable behaviour and the assumption that roles are unique to males and females (Harrison & Lynch, 2005). Every day, communities adapt to gender roles by gaining essential skills; for example, women and girls may gain domestic skills, such as learning how to cook. Such skills are inevitably gained because of the socialization of these positions. Social constructions are ensconced in such a way that people are required to perform certain positions, which is the foundation of many cultures (Koenig & Eagly, 2014). "To equip men and women for their usual family and employment roles, societies undertake extensive socialization to promote personality traits and skills that facilitate role performance. Role expectations thus exist in the minds of individuals and are shared with other people, producing the social consensus from which social structure and culture emerge" (Eagly & Wood, 2012, p.459).

Gender influences human behaviour. “As people view others in a particular social role, they then tend to generalize others who belong to that group as having the same characteristics, thus creating stereotypes” (Kiser, 2015, p.600). The norm includes information on the characteristics of the group. It postulates how individuals in that category behave, and if individuals fail to fulfil these criteria, they will be judged negatively (Tay et al., 2014).

Observations in paid jobs are another effect on gender stereotypes. Another point of interest is the division of gender into the occupation. The literature indicates that men hold a higher-paying role. This imbalance is a common belief that high paying positions need agentic attributes (Eagly & Wood, 2012; Koenig & Eagly, 2014). Gender roles are a central consideration of social dynamics in many cultures. They demonstrate how one gender should act and establish a common interest in what roles the gender should fill. At least for smooth social interaction, a person should conform and act according to given conduct to be consistent with, or avoid deviating from, a gender role (Williams et al., 2009).

Through these social beliefs about how women and men behave and how they should behave, children at a very early age learn about the roles of women and men as described by society and change their actions to conform and act accordingly (De Angeli & Brahnam, 2006; Mulvey & Killen, 2015). For example, a young girl’s behaviour and attitude towards football are affected by the first impression she has made, which is the lens she uses at that moment to decide if football is for boys, girls, or both. If she discovers that “rough play” is for boys, she is more likely to refrain from engaging in the sport than if she knows that every sport is for boys and girls (Mulvey & Killen, 2015). When children hear about these gender differences, they result in gender expectations that stick with them as they transition into adulthood. Such assumptions can be seen both at home and work (Heilman, 2001) and, when left unmanaged, contribute to disparity and lack of gender diversity, especially in managerial roles, as addressed in section 2.2 (Ridgeway, 2009).

Bian et al. (2017) said that 6-year-old girls have heard from cultural ideas that men are smarter than women. Such gender differences in intellectual capacity have led girls to assume less than boys that members of their same gender can be smart. They then shy away from the events that cultures set out for children with the intellectual ability to engage, such as playing video games. Such examples provide evidence that children learn these cultural concepts at an early age and that their actions filter what is understood about their gender.

Gender roles have an impact on one’s gender identity the sense of one’s self as a female or a male. “People use their gender roles as a standard guide to how they should act” (Eagly & Wood, 2012,

p.468). While agentic and social qualities are mostly expectations of society, people tend to internalize them (Williams et al., 2009). For example, many women identify themselves as emphatic, polite, supportive, and men as trustworthy and competent. "People vary in the degree to which they integrate gender stereotypes into their self-concepts. Individuals also vary in the types of gender roles that they follow. For example, men who regulate themselves as males may invest in masculine traits such as strength and competence" (Eagly & Wood, 2012, p.468). More details, specifically about social identity, are discussed in section 2.2.2. The critical presumption about societal norms is that people who are compliant with gender roles are likely to be rewarded rather than considered to be inconsistent. In effect, this continuity results in consistent interaction with those who meet social expectations and interrupt interaction with those who cannot (Ridgeway & Correll, 2004).

Most societies require women to show social qualities, and when women prove that they are agents, they will be penalized (Koenig, 2018). For example, since caring is often associated with women, if women do not support their co-workers, they may be penalized, while men may not. On the other hand, when men show supportive actions to their co-workers, they are likely to get rewarded than women (Bernhardt et al., 2018). As mentioned in section 2.2.3, male nurses are likely to be penalized; they are less tolerated than female nurses (Tay et al., 2014). Koenig (2018) states that, because these costs are known, men and women are less likely to violate the belief in gender roles unless the benefits outweigh the costs.

Eagly and Wood (2012) suggest that the theory of social roles provides an approach to understanding both continuity and transition in stereotyped occupations and associated gender disparities. The transition is evident as changes in women's roles have allowed women to access positions with a high degree of authority and income. It is a slight improvement. Research reports that there have been more women with masculine characteristics. Nevertheless, most men are likely to display communal attributes at home than at work. The transition is slower than that of women. However, there has been a rise in the contribution to childcare and some domestic work. Van Lange et al. (2011) states that women participate in childcare more than men because of this practice's continued importance for children. This system is fostered by female socialization and cultural values that encourage gender-specific role performance. Changes in gender equality are sluggish, often due to communities' philosophies and status values that allow for gender-based inequality (Eagly & Wood, 2012).

3.1.1. ORIGINS OF GENDER DIVISION OF LABOUR

Gender division of labour refers to how individuals divide jobs based on gender. This distinction applies not only to paid jobs but also to the distribution of duties and obligations at home between men and women (Eagly & Wood, 2012; Eagly et al., 2013). Social Role Theory claims that labour division induces gender differences in actions instead of the underlying psychological gender differences between men and women. The theory further suggests that the conditions contributing to labor division are distinct from the underlying physiological gender inequalities. The origins of the disparity in actions between men and women stem from the physical distinctions between these genders; they interfere with social and economic demands (Van Lange et al., 2011).

Physical gender discrepancies between men and women have contributed to the distribution of social roles in 'men' and 'women' positions, with men representing traits such as strength and women presenting warmth. Despite these physical distinctions, people assign those behaviours into one group instead of another, varying from culture. It has been noted that the gender division of labour is not 'one size fits all.' Societies have different perceptions of what is considered appropriate for both sexes. This is why, in some cultures, women perform roles that are performed by men in other cultures, such as being sole providers and vice versa (Koenig & Eagly, 2014).

Gender-based assumptions are derived from perceptions of men and women's position and represent the patriarchal division of labour (Eagly, 2013). Social Role Theory of Differences notes that "gender differences and behavioural similarities reflect gender role assumptions that, in turn, represent people's views of the social roles of men and women in the society in which they live" (Bobe & Kober, 2020, p.15). For example, men are more likely to be employed in practical roles, and women are more likely than men to be employed in caregiving roles (Eagly & Wood, 2012).

Often women's reproductive practices affect the occupation of certain professions. This makes it difficult for some women to participate as men. For example, when a woman is pregnant, she may not be able to be involved as men, particularly in activities that involve qualities such as uninterrupted facilities, speed, and being away from home (Eagly & Wood, 2012; Eagly & Wood, 2012; Koenig & Eagly, 2014). It is essential to note that not all women want to have children, and some work while the father does childcare (Eagly & Wood, 2012; Eagly & Wood, 2012; Koenig & Eagly, 2014). As a result, most women prefer tasks that can be done concurrently with child care. On the other hand, men's upper body, strength, and size often give them the benefit to occupy positions that require strength and size, such as the traditional position of hunting large animals (Eagly & Wood, 2012; Eagly & Wood, 2012; Koenig & Eagly, 2014). "Therefore, the division of labour

between women and men reflects the specialization of each sex in activities for which they are physically better suited under the circumstances presented by their society” (Van Lange et al., 2011, p.465). However, recently, these roles are less common in low-birth societies, where occupational roles, particularly in post-industrial societies, require less of these male-type attributes and differ across cultures. This change has resulted in incentives for women to have access to social roles that were initially masculine. (Eagly & Wood, 2012).

While there are improvements in the labor division, women still do domestic work more than men and spend more hours at home. Of most paid jobs, women's representation in the company's highest positions is low, maintaining a degree of patriarchy. (Koenig & Eagly, 2014). Gilligan and Richards (2008) state that patriarchy roots are defined when men and women's physical attributes with economic and technology development are attained more by men than women.

Through the lens of gender role beliefs from Social Role Theory, the researcher assumed that the participants would likely divide the bots' roles based on prior role belief assumptions. They are likely to associate midwife bots and feminine attributes to Pearl, and mechanic and masculine attributes to Peter.

3.2. SOCIAL IDENTITY THEORY

Both Social Role and Social Identity theories can be used to understand gender stereotypes from different perspectives (Guimond et al., 2013). Social Role Theory (SIT) mainly focuses on the continuity and transition of gender stereotypes (Eagly & Wood, 2016). It is mostly based on role beliefs. On the other hand, Social Identity explains how individuals categorize objects and define association and identities based on shared similarities, explained in detail by Similarity-Attraction Paradigm. The individuals then use the group's attributes they associate with to discriminate against other groups (Trepte & Loy, 2017).

The theory notes that people behave differently in various social situations according to the communities' norms to which they belong. It defines a person's sense of whom they are based on the group they identify with (Guan & So, 2016). In general, people believe that their group is better than other groups; this is to discriminate against their group and make them feel better about their group. By finding distinctiveness, people can also unconsciously perpetuate gender roles to boost their self-esteem. People may accept gender stereotypes about their groups to satisfy self-motivation by moving the blame away from themselves and towards the group (Trepte & Loy, 2017; Davis et al., 2019). For example, some people see mathematics as a topic where only men may study

and excel, so when a woman fails a math test, she is likely to support the assumption that “women are bad at math.” This acceptance protects her self-esteem from failure; however, it causes gender stereotypes to perpetuate in observers. This belief stems from the stereotypical norm that men excel well in mathematics compared to women. This belief may demotivate women and, in turn, impair their ability to excel in mathematics (Bian et al., 2017).

Social groups provide their members with a predetermined identity, which stipulates who they are, what they believe, and how they should act (Schmader & Block, 2015). Such categories are created by categorization, which is an essential part of the theory of social identity. People define groups by categorizing people based on common features, attitudes, and behaviours that illustrate similarities between inter-groups and differences between groups (Schmader & Block, 2015). Ridgeway (2009) states that categorization exists through identification. The identification of an individual is dependent on the discrepancies associated with another individual. Individuals are known as such since they are distinct from other people. When categorizing people into groups, individuals identify themselves as being connected, and this generates a common pattern that people use to categorize things and make sense of them.

People categorize objects in order to recognize and classify them. When individuals categorize groups, they create a prototype of the group. When a prototype matches their characteristics, they adopt that group's personality and behave in ways that the assumed members of that group behave (Treppe, 2006). Also, the Similarity-Attraction Paradigm (section 3.2.1) indicates that people are attracted to groups that are demographically like them (Wood & Eagly, 2015). The similarity in this sense refers to how individuals in a population are similar in terms of demographic characteristics such as age, ethnicity, and other behavioural traits (Pinar et al., 2014). From these differences, people identify and compare themselves with the outside community by seeking positive distinctiveness, which results in prejudice (De Angeli, 2005; Schmader & Block, 2015).

Such gender roles include the characteristics of a group, and people are generating standards about people belonging to that group based on these characteristics. Individuals who breach these standards appear to be viewed negatively (Tay et al., 2014). For example, since most people see nursing as a female profession, a male nurse may be viewed negatively.

Chaves and Gerosa (2009) contrast that by giving evidence to the benefits, these benefits include “increasing engagement, human-likeness, avoid negative stereotypes and balancing the identity and technical capabilities” (p.25-26). They state that the length of engagement increases as people ask questions about the chatbot's traits and status, although such chit-chats are often out of context to

what the bot is offering. Also, research shows that human-like attributes of the bot increase the naturalness of the chatbot, such as its name, the way it greets, and its language style (Chaves & Gerosa, 2009).

3.2.1. THE EFFECT OF ATTRACTION: SIMILARITY-ATTRACTION PARADIGM

In 1965, Bryne and Rhamey introduced a similarity-attraction theory, which stipulates that people appear to be drawn to those close to themselves. Early model descriptions named similarity as a positive relationship and attraction to the “law of attraction” (Ruijten, 2020). In 1975, Byrne argued that the paradigm is “one of the most robust relationships in all of the behavioural sciences” (Swami, 2015, p.162). Research has shown a positive correlation of interpersonal attributes with the number of similar attributes of others. Different studies have demonstrated this correlation that focused on attributes such as personality, background, and behaviour.

According to Social Role Theory, the primary roots of role beliefs are societal and cultural norms. The lens of Social Identity Theory and Similarity-Attraction Paradigm would also assume that the male participants will relate better to Peter and discriminate against Pearl and vice versa. The Social Identity Theory states that people are more likely to associate with those they share similar characteristics with than those with different characteristics (Bruner et al., 2017). Through this definition, the researcher related Social Identity Theory to Similarity-Attraction Paradigm, and this was adopted to explain how people associate and relate better to others who are similar to them. The Similarity-Attraction Paradigm defines this notion.

The principle of similarity-attraction has been widely used in psychology and communication studies. For examples, some studies are centred on interpersonal qualities (Montoya & Horton, 2004; Schug et al., 2009); on relationships (Morry, 2005; Tidwell et al., 2013), and the repulsion hypothesis that is contrary to similarity-attraction (Chen & Kenrick, 2002; Singh & Ho, 2000). Rosenbaum's repulsion theory notes that it is not by similar attitudes that people are drawn to each other, but instead, people repel when they learn different attitudes from others (Chen & Kenrick, 2002). These similarities are likely to give rise to the initial attraction between people.

The Similarity-Attraction Paradigm is also evident in the online world from a human-to-technology interaction. Most of the research from this perspective focuses on the same-personality interaction between humans and the bots such that introverted humans communicating with introverted bots, and vice versa (Bernier & Scassellati, 2010). Other studies looked at different interaction styles, such as the submissive vs. persuasive interaction (Ruijten, 2020). Bernier and Scassellati (2010) argue that

contact between humans and embodied agents has shown that individuals prefer bots with similar personalities. In their study, participants rated robots that shared similar attributes as friendlier than robots with different attributes. On the other hand, Lee et al., (2006) study found that “extroverted participants viewed the introverted robot as more intelligent and socially desirable than the extroverted robot, whereas introverted participants favoured the extroverted robot” (p.287).

Previously, the similarity-attraction was contested as there was no explanation for presenting empirical evidence. One of the reasons is the inherent inclination of individuals to conform to the standards of society. This explanation assumes that individuals’ characteristics similar to a group confirm their sense of belonging and views. It also led to the development of new models to explain the Similarity-Attraction Paradigm. There are models such as the repulsion hypothesis by Rosenbaum, which did not gain much attention. The reinforcement model and the information perspective have received both support and criticism for being able/unable to provide empirical evidence (Kaptein et al., 2014). For example, Byrne’s work could not explain why attraction occurs more in the laboratory and not field studies.

Montoya and Horton (2013) state that the model failed to clarify the attraction of negative traits between people. Some researchers challenged the information perspective on why the similarity of less significant attributes does not contribute to less attraction than similarity to critical attributes. It also does not affect the number of attitudes that individuals consider as similar to those of others.

These theories are in opposition, and the literature is not clear whether the influence of the Social Role Theory, Social Identity Theory, or Similarity-Attraction Paradigm is stronger than the other.

3.3. SUMMARY

The below section explains how each Social Role, Social Identity Theories, and Similarity-Attraction Paradigm were relevant to answer the research questions.

The theories have enabled the researcher to see how gender differences and similarities in human-to-chatbot interactions influence people's behaviour. Social assumptions about gender are at the center of these attitudes, as clearly stipulated in the Social Role Theory. These stereotyping, or gender role assumptions, form when people perceive male and female behavior and infer that they are different. Prior understanding of gender roles and the gendered division of labor also affects interactions. As a result, the researcher addressed questions about the effect of gender on participants' feelings, reactions, beliefs, and expectations. The researcher noted that even though society has developed views about gender roles, some members have views contrary to society.

Following these trends from the Social Role Theory, the researcher addressed how people respond emotionally to chatbots of different genders to see whether there is a role of belief in this case. This theory was useful in answering the research question about participants' feelings and how they reacted to chatbots of different genders, how they communicated with each chatbot, and how the gender of the chatbot in the roles met their expectations.

Social Identity Theory and Similarity-Attraction Paradigm helped address the research questions from a different viewpoint than Social Role Theory. Social Identity Theory and Social-Attraction Paradigm in this research were more evident during the interaction, and there was no evidence that the prior-knowledge of gender roles influenced them. The reason may be that similarities between the participants and the bots were evident during the interaction. Once these similarities were observed, the participants related to the bots and gender stereotypes prevailed. Unlike in Social Role Theory, where societal norms formulate gender stereotypes, the Social Identity Theory and Similarity-Attraction Paradigm added to this theory as some of the gender stereotypes remarks from the participants were because they share similar characteristics with the bots. For example, a female participant preferring a female bot over the male bot. There was more than one characteristic of the bot in which the participants could have identified the similarities, such as voice or the name. However, they related to bots based on gender. This, in turn, stipulated that Social Identity Theory and Similarity-Attraction Paradigm were also applicable to address research questions about how people respond emotionally to chatbots with which they share similarities, in this case, gender. These similarities and differences in gender have, in turn, been able to address questions about the influence of the gender of chatbots on user responses and experience.

4. METHOD

This study aimed to explore the phenomenon of chatbot's gender stereotypes in South Africa using chatbots of different genders performing the same role with the same content to raise awareness of gender stereotypes in chatbots' design.

For this research, the epistemology that was adopted was interpretivism. The researcher's philosophical assumptions guide the decision about knowledge and how to obtain it. The assumption was that the reality concerning individuals' perception of chatbots' gender in South Africa could be obtained through social constructions. "The core idea of interpretivism is to work with these subjective meanings already there in the social world; i.e., to acknowledge their existence, to reconstruct them, to understand them, to avoid distorting them, to use them as building blocks in theorizing" (Goldkuhl, 2012, p.5). This is a qualitative study that provides an individual's perceptions of the phenomenon. It provides insights into individuals' perceptions of the gender of chatbots.

The research approach that was adopted was 'partly inductive.' The research approach was not purely inductive since Social Role Theory, Social Identity Theory, and Similarity-Attraction Paradigm were adopted as guidelines to collect and analyse data. The literature review was also conducted, which shed some light on the research questions' expected answers. Likewise, the approach was not purely deductive since there were no hypotheses tested. The deductive aspect, which emerged in using the theories, enabled the researcher to elicit data that answered the research questions. Chatbot's gender stereotype is a new phenomenon, and this has led to the emergence of new concepts, as outlined in section 2. An inductive approach gave the researcher the ability to consider the emerging concepts, which added to the body of research regarding this phenomenon. Through an inductive approach, the researcher identified patterns and relationships from the data to build upon the used theories to provide a different approach to understanding the phenomenon. There is a lack of theories to study individual perceptions of chatbots' gender in the Information Systems field.

4.1. DATA COLLECTION STRATEGY

Data was collected qualitatively through chatbot interaction and interviews (Robinson & Mendelson, 2012). Although many researchers have collected data qualitatively, few have created an experience for participants, followed by interviews (Robinson & Mendelson, 2012). Collecting data

through chatbots and interviews did not only result in in-depth data but also ensured the study's validity.

The researcher created an experience for the participants by developing chatbots to determine the participants' perceptions when interacting with chatbots of different genders and the same function. It also measured how the participants responded emotionally to chatbots of different gender performing the same role in a controlled environment (RQ 1.1). Following this, interviews were conducted to extract in-depth insights into the phenomenon while also matching the interviews' outputs with that collected after participants' interaction with the chatbots to answer RQ 1.2-1.4.

The type of interview that was adopted was semi-structured interviews, which allowed the researcher to get more information on insights that emerged for a richer understanding of the phenomenon. The interviews were also used to probe and explain why there is or is not a cross-gender preference in this study, which, in turn, contributed to the methodology as there is little evidence of this preference in the body of knowledge. The emergent of this preference is discussed in sections 6 and 7. The methodological contribution is from using experience with interviews in qualitative research. Triangulating data collection strategy ensured that the data was rich and in-depth (Fusch & Ness, 2015).

In phase I, the participants were taken through experience by interacting with the chatbots, followed by phase II interviews. Both the phases were conducted on the same day per participant at the University of Witwatersrand. Each participant was given a slot as per their advised availability to interact with the chatbots and then interviewed. The researcher was there to explain the procedure to the participant before phase I; however, the participant was alone when interacting with the chatbots. Following the interaction, the researcher joined the participant for the interview. The discussion was recorded, and the researcher referred to these recordings and analysed the responses at a later stage, as discussed under the findings to answer the research questions. The time horizon for this research was cross-sectional because the purpose of the research was to explore individuals' perceptions of the gender of chatbots in South Africa at a single point in time. The aim of the research was not to investigate changes to perceptions over time.

4.2. DATA COLLECTION INSTRUMENTS AND DESIGN

This section discusses the tools used to collect data and details the instruments' design adopted. Pilot testing was done both for the interaction with the chatbots and interview questions to identify any discrepancy in the instruments and make sure that the prospective participants understood the process. This ensured that the instruments were designed to collect adequate data that successfully answered the research questions. The pilot testing was done with six individuals from Standard Bank South African between the 15th and 19th of August 2019.

4.2.1. EXPERIENCE DESIGN

The experience was based on a 2 x chatbot gender (male vs. female), 2 x participant gender (male vs. female), and 2 x stereotyped subject domain (auto mechanic and midwife) matrix, as demonstrated in Table 1 and Figure 2. Literature stipulates that these roles are stereotyped (Baker et al., 2008; Clow et al., 2014). According to the literature, practical roles are perceived as males and caretaking or nurturing roles for females (Knobloch-Westerwick & Hoplamazian, 2012). For this study, two gender-stereotyped roles were used. In this case, the assumption was that auto mechanic service providers are perceived to be males and midwives to be females.

Section 2.2 stipulates gender stereotypes in the nursing field, which are assumed to be like those of midwives as they share similar responsibilities specifically for childbirth and care. For example, a male chatbot offering midwife services may be stereotyped because it does not conform to gender role beliefs as stipulated in the literature. The chatbots were developed from scratch. There are pre-built chatbots that could have been adopted to give participants this experience; however, their roles were not suitable for this research.

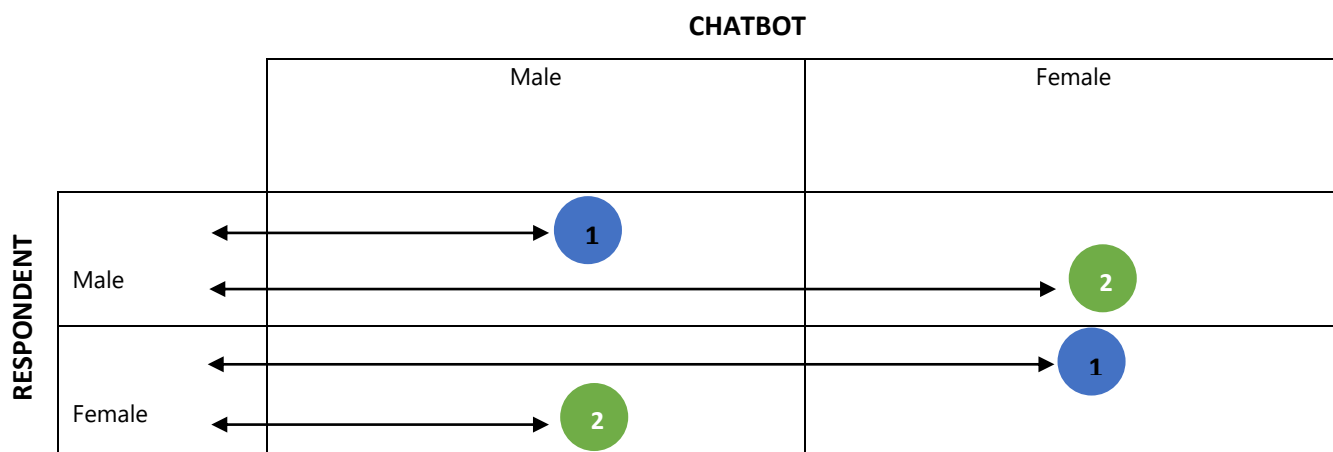


Table 1:High-level Interaction Process

← → The arrow represents the interaction between the gender of the participant and that of the chatbot.

1 Interaction one represents an interaction between the participant and the chatbot of the same gender, for example, a male participant interacting with a male chatbot.

2 Interaction two represents an interaction between the participant and the chatbot of a different gender, for example, a female participant interacting with a male chatbot.

Each participant engaged with both the male and female versions of chatbots of one role. The first half (4 male and female participants) interacted with the chatbots offering auto mechanic services and the other half with the chatbots offering midwife services:

Let the first four male participants = **a**

Let the first four female participants = **b**

Let last four male participants = **c**

Let last four female participants = **d**

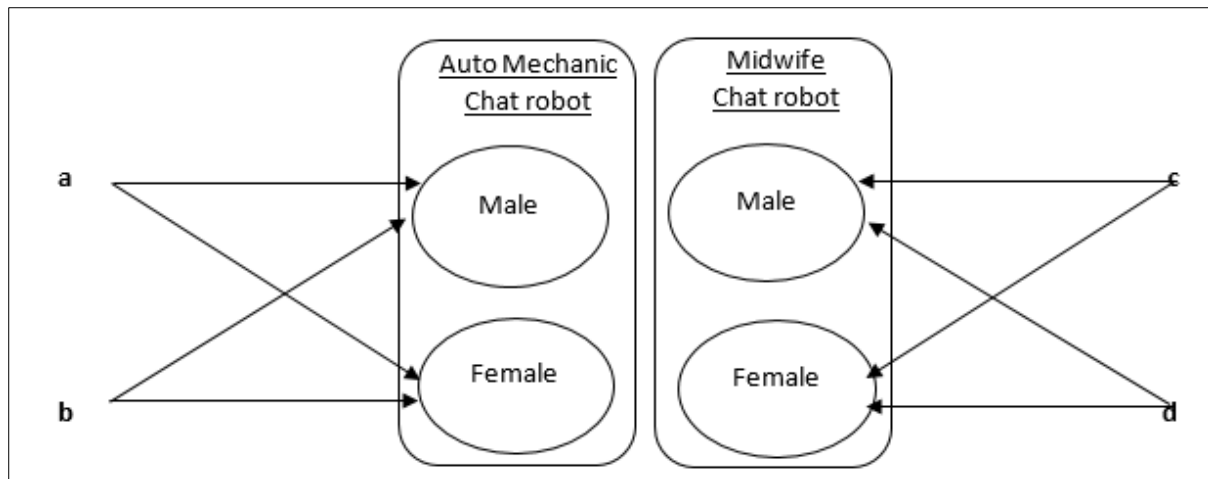


Figure 2: The interaction of participants with both the male and female version of one of the roles.

Initially, the participants would interact with the bots of the unexpected gender in the stereotyped roles. For example, for the mechanic bots, the participants would have interacted with a female bot, and for the midwife bots, the male bot would have been the first encounter. During the interviews of the first two participants, comments about gender did not emerge much. This has left the researcher concerned about whether the instruments were designed to collect data to answer the research questions successfully. The researcher decided to swap the genders for half of the participants to test if the responses were the same regardless of the gender of the chatbot, as demonstrated in Table 2 below. The effects of stratifying the experience in this manner are discussed in detail in section 6.5.

Participant	Auto-mechanic chatbots		Midwife chatbots	
Gender	Pearl-Peter	Peter-Pearl	Pearl-Peter	Peter-Pearl
Females	2	2	2	2
Males	2	2	2	2

Table 2: Actual Interaction Process

Table 2 demonstrates the actual interaction process between the participants and the chatbots. For each group, two female and male participants interacted with Pearl first, then Peter and vice versa.

The auto mechanic chatbot provided car diagnosis services for car repairs. For the scope of this research, the service was only for a car that does not start. Similarly, the midwife chatbot service was for contraceptive related inquiries.

The participants were given one of the two below scenarios to start a conversation with the chatbots. Participants who interacted with the auto mechanic chatbot were requested to require assistance from the chatbot to repair a car that does not start. Also, the participants who interacted with the midwife bots were requested to require assistance with contraceptives. The scenarios were used to narrow the scope of the chatbots' development and interaction. Also, to ensure that the responses were drawn from the same context.

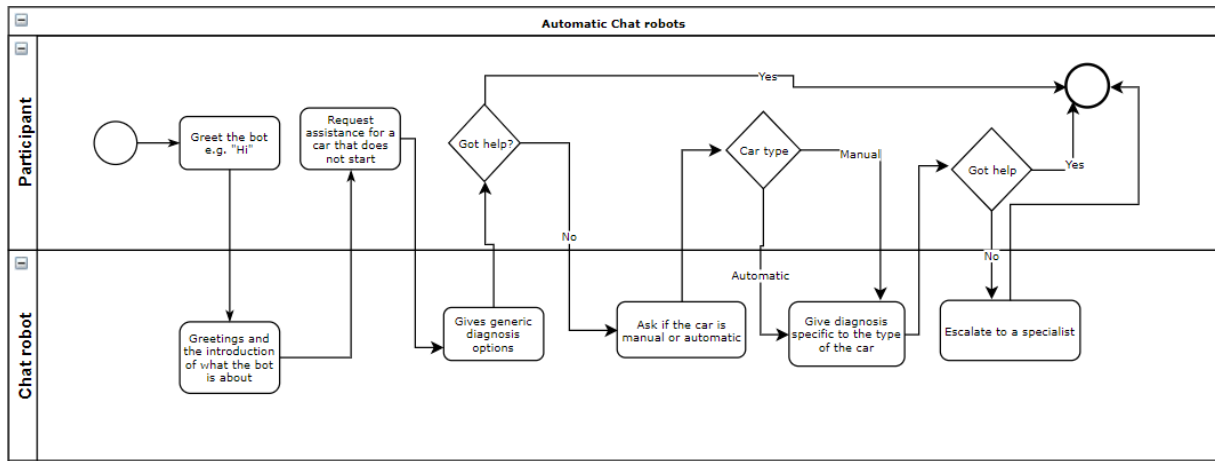
Scenario 1: "It is a Saturday morning, and you just woke up to make breakfast. You realize that there is no milk in the fridge. You quickly grab your car keys to buy some milk. You then find out that your car does not start from the last time you have parked it in your garage, and you do not know what to do. Use the auto mechanic chatbot to ask for assistance, assuming that you have no idea about cars."

Scenario 2: "You are looking to get into a full-term relationship commitment, and you have no idea how to avoid unwanted births until you are ready for children. Use the midwife chatbot for assistance, assuming that you have no idea about contraceptives."

Further instructions were given to the participants to ensure an uninterrupted experience. The participants were asked the following:

- Do not close or refresh the tab. To restart the conversation, greet the bot. This was done to avoid the loss of the interaction between the participant and the bots.
 - Try to stay on the chat as long as possible. Do not talk to the bot to break it.
- Please check your spelling because of the bot could not correct all the incorrectly spelled words.

Figure 3 depicts the generic interactions between the chatbot and the participants based on how the bots were developed.



47

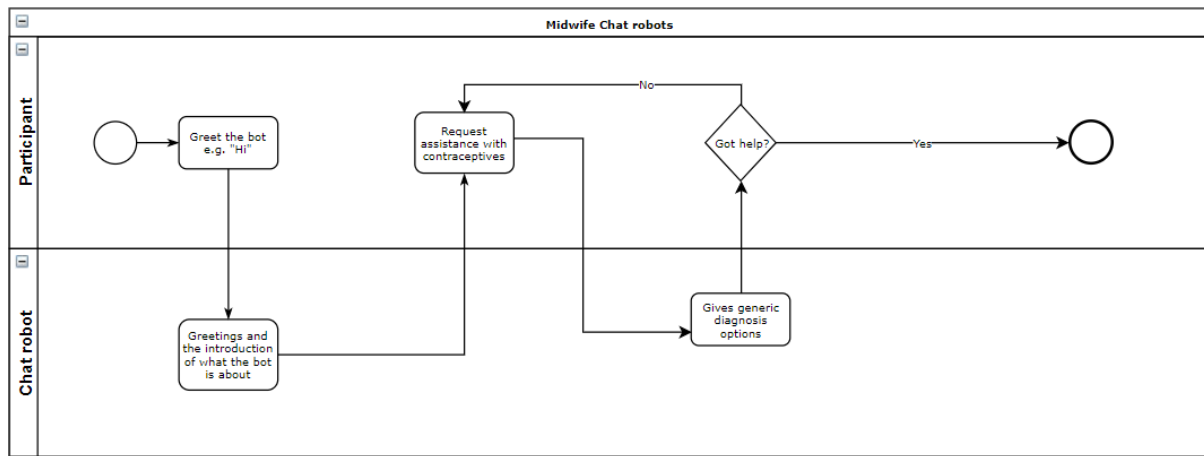


Figure 4: Engagement flow between the participant and a midwife chatbot.

Figure 4 provides a high-level example of the interaction between participants and the midwife service chatbot. Participants were asked to greet the bot in order to launch the conversation. The chatbot then presented itself by presenting information about its name and the services it offers. An example of the request was, “Hello! What is family planning?”. The chatbot then displayed the answer to the question. Unless the participants have more questions to ask, submitting a question will be repeated before they feel that he/she has received help. If the bot did not help, the inquiry was escalated to a specialist; then, the chat was ended. If the bot helped, the talk would be over.

4.2.1.2. CHATBOT FEATURES

Siegel et al. (2009) suggested that future research looks at other gender components than the voice changes because a chatbot with only verbal and text responses is not sufficient for more significant outcomes regarding perceptions. A combination of physical appearance has more effect on an individual (Baylor, 2009). Baylor (2009) reflected that students who interacted with visible chatbots with embedded physical traits such as an avatar showing facial expressions reported a significantly more significant effect. However, the physical appearances were not considered for this study because the researcher wanted to explore the phenomenon without designing chatbots that could potentially lead to obvious gender identification and differences. Radziwill and Benton (2017) defined conversational agents with physical appearance as embodied conversational agents and not chatbots. Therefore, for this research, the selected components were the voice and the typical male (Peter) and female (Pearl) names. The chatbots responded differently, only based on the role. For example, a chatbot offering auto mechanic services responded differently to the one offering midwife services. However, the responses of the chatbots of the same role were identical.

Figure 5 is an example of conversation architecture. It stipulates that users were able to communicate with the chatbot either by text or voice prompts. If the user conversed through voice, the voice was converted to text. This capability is known as speech-to-text. Likewise, the chatbot responded in text and voice only.

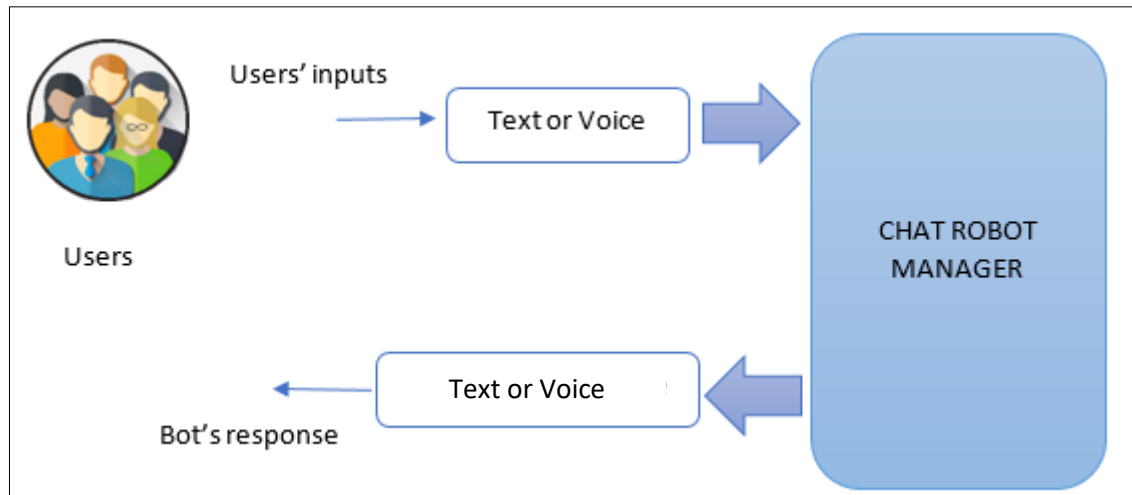


Figure 5: Basic architecture of the conversation between users and the chatbot.

Figure 6 stipulates the architecture of chatbots' design in Dialogflow, a platform used to develop chatbots for the experience. Similarly to the architecture in Figure 5, a user-triggered the chatbot through text or voice. Only two participants mentioned that they tried using the voice capability to communicate to the bots; however, that was not successful because they translated the voice to text incorrectly. One possibility is that this could be due to the challenge of understanding the accent of the participants' voices. The input was sent to the chatbot by Dialogflow. Intents are defined in Dialogflow to categorize the inputs. They define the intended goal of the user. For example, how to get to Johannesburg can be categorized into an intent called directions. Intent does not respond to the input but provides the bot with enough context to make a decision. The chatbot then responded to the input through text and voice.

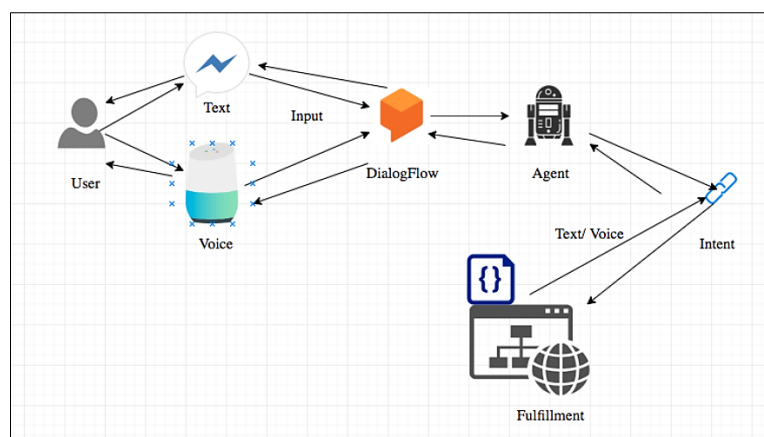


Figure 6: Flow of conversation within DialogFlow.

4.2.1.3. TOOLS FOR CHATBOTS DEVELOPMENT

The chatbots were built on Dialogflow, Google's online platform, and were incorporated into and accessible via the Dialogflow web portal. Dialogflow allowed the developer to pick a customized voice for the chatbot to transcribe text content to audio, and the default accent was American. The voices of the chatbots and the accent were adopted as they were configured on the website. The chatbots have been reached via the Dialogflow web interface. It was easier to integrate the Chatbot on the channel, as Dialogflow provides this integration with a few clicks, and no coding was needed. The participant had to start the conversation by greeting the bot.

To document the communication between the participants and the chatbots, screenshots were taken and saved to record the conversation after each participant's contact. Such recorded conversations were preserved and then reviewed, along with the transcripts of the interviews. Figure 7 is an example of a chatbot deployed on the Dialogflow and ready to be accessed.

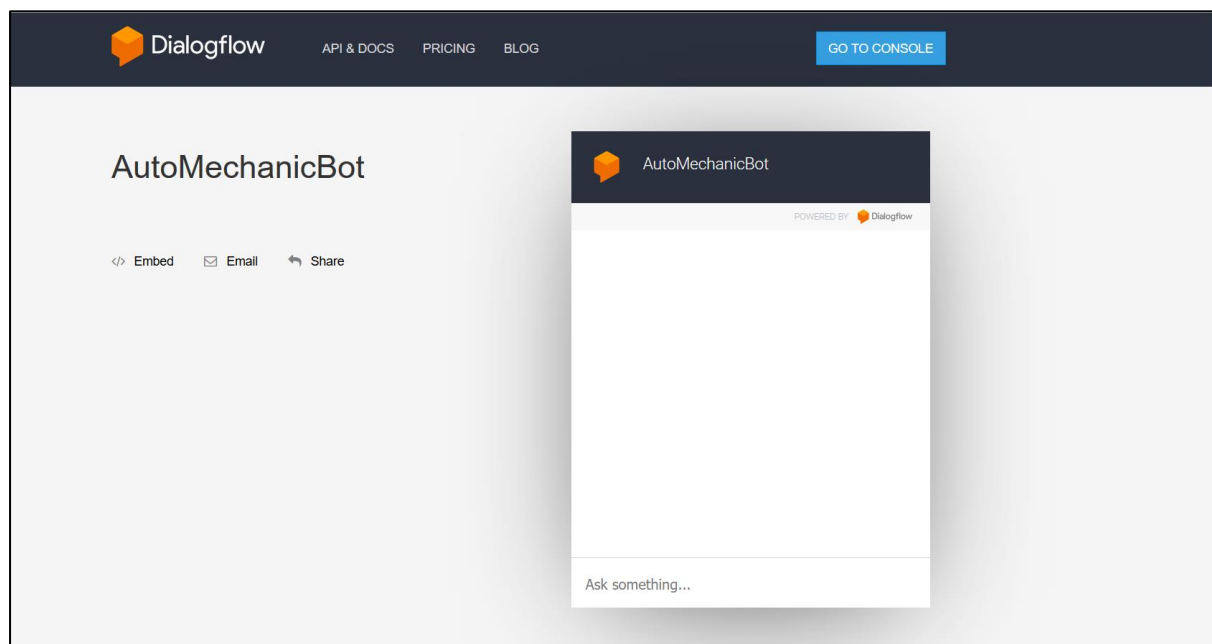


Figure 7: A chatbot deployed on Dialogflow Web.

4.2.2. INTERVIEW DESIGN

Interview questions were extracted from the theoretical frameworks adopted for this study, as set out in Table 3. Interview questions on emotions, grouped under RQ1, were guided by all the frameworks. The presumption was that applying the Social Role Theory to these issues, how participants respond emotionally to the chatbots may be influenced by the gender of the chatbot. This meant that participants would emotionally react positively towards the chatbots that meet their expectations and vice versa. Also, arguments from the Social Identity Theory and the Similarity-Attraction Paradigm provide evidence that it is not only prior knowledge of gender roles that influences relationships but also mutual similarities. This meant that while some participants' emotions are influenced by the socialization of gender roles from their cultures, some are influenced by mutual characteristics with chatbots. This meant that the participant would respond positively to the chatbot if it were of the same gender.

The literature states that how some people interact is influenced by gender. The interview questions about the chatbots' gender influence on participants' responses and experience were derived from the claims made by all the frameworks. Applying these frameworks from this perspective meant that we could have insights into the impact of the chatbots' gender on users' interactions. Social Identity Theory and Similarity Paradigm stipulate that there would be a positive interaction between participants sharing the same gender with the chatbots than when they are not. In this instance, Social Identity gives a different view concerning this issue. It stipulates that the chatbots' femininity and masculinity concerning role stereotypes affect how they interact with the chatbots.

In this sense, the prior-existing expectations are social norms that equate nursing with women and mechanics with men. The Social Role Theory offers evidence that the gender division of labor contributes to these assumptions, suggesting that the participants will then have assumptions of the chatbots' characteristics performing that role when provided with the chatbots' roles.

The interviews were semi-structured, as detailed in section 4.1. They were carried on the same day with the experience as phase II of the data collection strategy. A digital recorder was used to record the interviews, and permission from the participants was requested before the recording. The recordings were kept anonymous and will be destroyed after the submission of the paper. The researcher probed for gender throughout the interview if it did not emerge, while not negating any other reasons for differences in responses and experiences. The interview responses were transcribed using an online tool, Sonix. After transcribing the responses online, the researcher exported transcribed transcripts to a word document to check for correctness. The researcher listened to the

recording alongside the transcripts to check if the interviews were transcribed correctly by Sonix and made necessary corrections. The interviews were stored in a password protected OneDrive account. The following table detail the interview questions that were asked.

	General Question	<ol style="list-style-type: none"> 1. Please tell me your age range 2. What gender do you identify as? 3. Where do you stay? 4. Do you have a smartphone? 5. Please describe your use of social media, if any 6. Please describe your experiences during any previous use of chatbots
Theoretical Mapping	Research Question	Interview Questions
Social Role, Social Identity Theories, and Similarity-Attraction Paradigm	RQ1: How do people respond emotionally in a controlled environment to chatbots of different genders performing the same role?	<ol style="list-style-type: none"> 7. Please describe if you felt there was a difference for you in being assisted by Peter or Pearl 8. Please describe your feelings when Peter offered you midwife services as opposed to Pearl OR Please describe your feelings when Pearl offered you auto mechanic services as opposed to Peter
Social Role, Social Identity Theories, and Similarity-Attraction Paradigm	RQ2: What is the influence of a chatbot's gender on the user's responses during the interaction?	<ol style="list-style-type: none"> 9. Please explain if your responses were the same for both the chatbots and why or why not 10. Please explain which chatbot you felt more comfortable interacting with and why 11. What influence did each chatbot have on your responses?

Social Role, Social Identity Theories, and Similarity- Attraction Paradigm	RQ3: What is the influence of a chatbot's gender on user's experiences?	<p>12. Please describe whether you believe Pearl and Peter are the same in terms of service delivery</p> <p>13. Please detail what characteristics would you change about the chatbots and why</p>
Social Role Theory	RQ4: How were people's prior-existing expectations of the role of the chatbot affected by its gender?	<p>14. Please describe if you felt that the chatbots were skilled for the role?</p> <p>15. Please explain if Peter and Pearl met your expectations and why</p> <p>16. Please explain whether there is any gender bias regarding the role of auto mechanics/midwives in your culture or society</p> <p>17. For this service, would you have a preference for either Peter or Pearl? Please explain your answer.</p>

Table 3: Interview questions.

4.3. SAMPLING AND SAMPLE SIZE

This section discusses the researcher's process to sample the participants and explain the chosen sample size.

Under the non-probability method, various techniques can be selected for sampling, such as convenience, quota, judgment, or purposive (Etikan & Bala, 2017). The technique that was used is the convenience sampling technique and theoretical sampling. In a convenience sampling, the participants selected to be included in the sample are the easiest to access (Acharya et al., 2013; Etikan et al., 2016). For this research, there was no equal chance that the individuals will be selected as a sample. The participants were selected based on willingness and availability to take part in the study. Also, it is stated from the literature that the non-probability sampling method is mostly suited for exploratory studies. The theoretical sampling selected eight female and male participants for equal representation of each gender, allowing the researcher to analyse chatbot stereotypes across both genders. The summary in section 6.4 shows the difference between female and male participants' perceptions of gender stereotypes.

Social media platforms, Facebook, WhatsApp, and LinkedIn, were used to reach out to the participants, and participation was based on the participants' availability and willingness. The researcher posted the invitation for participation on social media platforms and asked the prospective participants to send a direct message if they were willing to participate. Social media platforms have enabled people to send private messages using applications such as Messenger for Facebook. The information about the date and time of the experience and interview was communicated directly to each participant through social media; the schedule is detailed in annexure 1. The assumption was that participants on social media would be more computer literate, therefore completing the experience. There were sixteen (eight males and eight females) participants selected to interact with the chatbots and be interviewed. Gender was controlled for the chatbots to align with the theory and methodology adopted for this research and fulfill the purpose of creating an experience for the participants. Gender differences and similarities, as defined by the theory and literature, were explored through participants interacting with the chatbot of the same gender and opposite gender.

The interview sample size was selected to achieve saturation, which defines the data collection point, where no new themes and insights emerge (Fusch & Ness, 2015). Guest, Bruce, and Johnson (2006) found that even though they conducted 60 interviews for their study, they reached a saturation point after twelve interviews, with most of the themes emergent after just six. Fusch and

Ness (2015) stipulated that there is no one size for everything; more participants are not necessarily better than fewer participants and vice versa. They have declared that data triangulation results in thick and rich data, resulting in data saturation. “There is a direct link between data triangulation and data saturation; the one (data triangulation) ensures the other (data saturation). In other words, data triangulation is a method to get to data saturation” (Fusch & Ness, 2015, p.1411). They have argued that not all data methods result in data saturation. However, they have given evidence that interviews, focus groups, and data triangulation are often resultant in data saturation amongst all data methods.

5. DATA ANALYSIS PROCESS

This section describes the approach used to interpret the data and explains the steps taken to do this. The method of analysis of data that was introduced was thematic analysis to obtain in-depth insights. Most researchers in Information Systems use thematic analysis to interpret qualitative studies (Nowell et al., 2017). It allows the researcher to minimally explain the data set in more in-depth detail by grouping data into themes. “The theme captures something important about the research query data and reflects some form of patterned response or context within the data collection” (Braun & Clarke 2006, p.10). A list of themes was derived from the transcribed data.

Data were obtained and qualitatively analysed to align the interviews' findings with the findings of the experience. The researcher preliminary analysed the data followed by further in-depth analysis of the whole data set. Analysing the data for each participant helped the researcher recognize possible gaps that required a different approach to produce more in-depth insights to address the research questions; thus, the researcher could swap the gender of the chatbots as outlined in section 6.5. The researcher obtained the data in conjunction with the analysis. There is an “inseparable relationship between data collection and data analysis, which is one of the major features that distinguish qualitative research from traditional research” (de Vos et al., 2005, p.335).

When interpreting the data, part of the process included the researcher's interpretation of how to make sense of the data obtained. Interaction with each participant during the interviews provided the researcher with the benefit of obtaining previous knowledge for the next interview. This process made it possible for the researcher to come up with inductive themes emerging from the data. The research method involved transforming data into conclusions by ordering, structuring, and adding sense to the data collected. The researcher did not pursue a linear method of thematic analysis; the phases were implemented iteratively. This approach includes reducing the amount of information, recognizing trends, and categorizing related content (Neuman, 2014). Data collection and interpretation were closely intertwined. The researcher analysed and interpreted data in parallel.

5.1. GENERATING THEMES

The challenge with thematic analysis is understanding each data set from all participants and ensuring that all facts are preserved and represented. In achieving this, the researcher regularly went through the transcripts to understand what each participant was communicating before analysis. The understanding produced from the preliminary analysis of one participant made it possible for the researcher to understand what the other participants conveyed. Conversely, the theme analysis also involved noticing how one participant's responses fit into the chosen theme, while another may have indicated a deviation from the theme.

The researcher used the coding process to organize and establish order from the various patterns discovered.

5.1.1. CODING OF THEMES

The codes were captured online on the Atlas.ti cloud. Atlas.ti launched the cloud beta version, which is available to the public for free until they move it to alpha. The codes were added to the transcripts, uploaded, and grouped to shape overarching themes, contributing to the theory's developments.

Initially, the researcher established the initial codes from the first 10 participants, including theoretical and literature concepts detailed in section 1.1.2. The researcher went through the other transcripts to find repeated patterns, break them down into smaller coherent terms, and carefully analyse them for variations and similarities. Throughout this step, the researcher used the colour coding given by Atlas.ti to mark the codes into similar themes that are most logically related to the data, mainly for the researcher's readability. Also, the codes were divided into groups based on shared meanings. This feature was commonly used by the researcher to draw data from codes grouped under a common theme. For example, in Figure 8, quotes with evidence of feelings around comfortable, impersonal, and frustrations were group under emotions.

The researcher then concentrated on the individual data and allocated codes to themes to identify the relationship between themes. Different themes appeared, and some of the themes made sense to be grouped because they seemed to be distinct at first, but as the researcher drew more attention to themes, they reflected the same concepts. For example, the theme "user responses" and "influence on user responses" were combined into the theme responses. User responses initially consolidated responses to determine if the participants' responses to the bot were the same or not.

The influence on the user response theme is determined if factors such as gender influenced how the participants responded to the bots' questions. During the analysis, the themes related such that the participants' responses were the same or not for the bots because of certain factors that influenced their interactions as detailed in section 6. The challenge was to deal with converging and divergent themes without sacrificing the concepts' clarity and sense. To accomplish this, the researcher often asked questions such as, "can I divide established concepts into subcategories?" and "can I bring these ideas together?" Some codes from the theory and literature, such as gender hierarchy, were removed because they had no connection with the transcripts.

Finally, the researcher then reorganized the themes into small manageable themes for final representation. This was after all the data had been gathered. The goal was to search all data and previous codes for selective codes to be core categories in which other codes were grouped.

5.2. THEMES

This section summarizes the developed themes before data collection and the final themes after data collection, as shown in Figure 8. Only the gender hierarchy was removed from the initial themes because there was no evidence of this theme in the data. Some of the themes are re-grouped and renamed to make sense of the data (see section 6 for more details).

The themes before data collection were:

1. Division of labour
2. Gender stereotypes
3. Gender role beliefs
 - Violation of gender role beliefs
4. Society and culture - Socialization
5. Gender hierarchy
6. Gender-based characteristics
7. Service evaluation

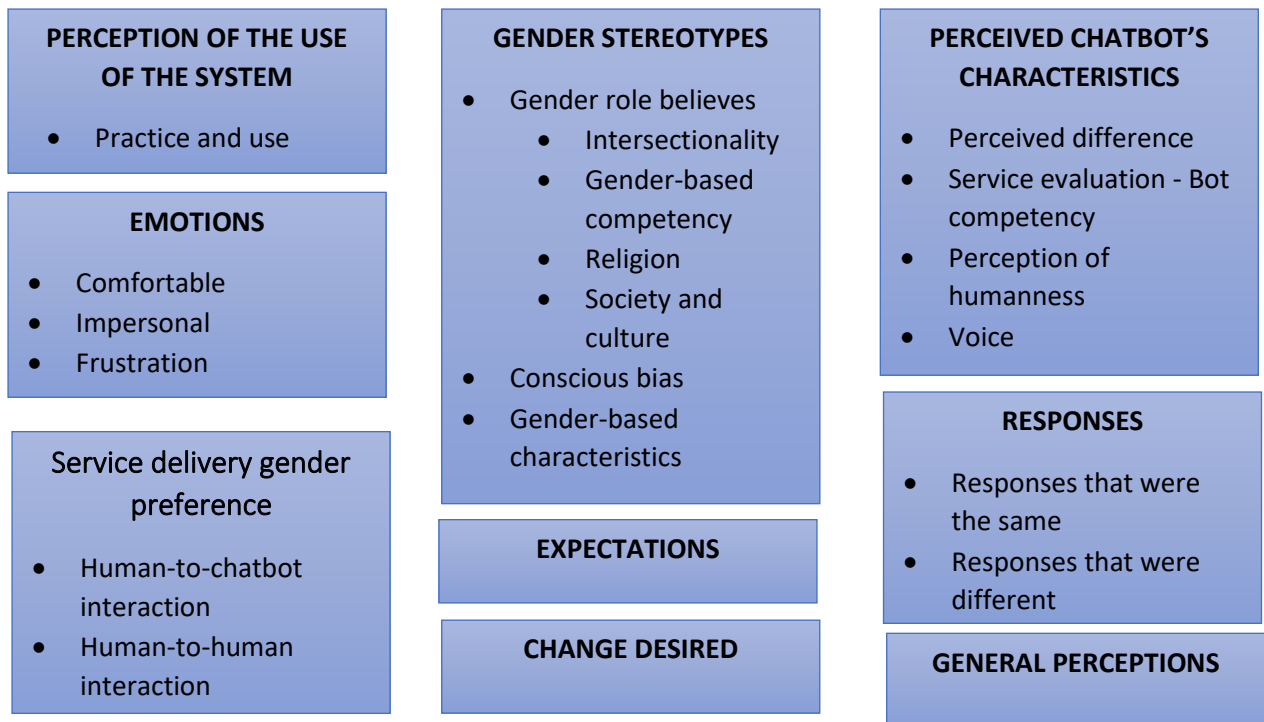


Figure 8: Themes after data collection.

The interpretation of data following Neuman's (2014) data analysis process allowed the researcher to develop themes that helped to answer the research questions successfully. The following section, section 7, consolidates the responses into themes.

6. RESULTS ANALYSIS

The participant went through the experience and interviews on the same day of their participation. The experience was based on a 2 x chatbot gender (male vs. female), 2 x participant gender (male vs. female), and 2 x chatbots' roles (an auto mechanic and midwife agents) matrix. This section consolidates the analysis and findings from the participants' responses obtained by interacting with the bots and interviews. It relates the findings to previous research and demonstrates the relationship between themes. It is divided into different themes and sub-themes labelled in section 5.

6.1. DESCRIPTIVE ANALYSIS

Sixteen participants participated in the research and went through both the experience and the interview processes. Out of the sixteen, eight (four females and males) interacted with both male and female chatbots offering midwife services, while the other half interacted with chatbots offering mechanic services. Eight of the total participants described themselves as females and eight others as males. All participants reside in Gauteng and differ in age, as outlined in Table 4. They all have smartphones and have a presence in social media for several purposes. Most of them use social media to socialize, keep up-to-date with what is going on worldwide, personal purposes such as fitness inspiration, and improve their companies. Eight participants had never communicated with chatbots before—the other eight communicated with chatbots via media such as Siri, AliExpress, and Hackathon.

Participant Pseudonym	Gender	Age-range	Residence	Own a smart phone ?	Use of social media	Previous experience with chatbots
P1	Female	20-30	Centurion	Yes	"High, I would describe my use of social media as very high."	Closest experience is with SIRI
P2	Male	20-30	East rand	Yes	"It is more business and interacting with people for business."	Have created one during a hackathon
P3	Male	40-50	Johannesburg	Yes	"I follow mostly news sites and personal interests. I do	None

					not necessarily post to a lot myself or interact in terms of feeling that I should necessarily contribute to every social discussion that is happening out there.”	
P4	Female	20-30	Johannesburg	Yes	“I use social media [...] just for general purposes, posting pictures.”	None
P5	Male	20-30	Tsakane, East of Johannesburg	Yes	“[...] for socializing with people’ and just being updated on the news around the world since when I do not have T.V. and for gossip”	None
P6	Male	20-30	Carletonville	Yes	“[...] to socialize.”	None
P7	Female	20-30	Johannesburg	Yes	“I use Facebook for stalking people. Okay.... check on what other people are doing. [...]. No, I have Facebook. I go to Instagram to get fitness motivation. Okay, and then Facebook is more like memes, just like Twitter.”	Closest experience is with AliExpress shops
P8	Female	20-30	Germiston	Yes	“I use a lot of social media, but specifically, it is WhatsApp and Facebook if I was to rank it WhatsApp, Facebook. [I use social media] for socializing mostly.”	None

P9	Female	20-30	Pretoria	Yes	"For social stuff, personal stuff."	Used one from the travel booking sites
P10	Female	30-40	Soweto	Yes	"Facebook, I use it to post photos and update status. For socializing"	None
P11	Female	20-30	North Cliff	Yes	"If any [I] use social media for business and private use."	None
P12	Male	20-30	South of Johannesburg	Yes	"I touch my phone every 10 minutes unless I am really busy, then I would leave it alone for a good 30 minutes. Instagram is the best one, WhatsApp I have Facebook, but I do not have the app. I also have a Twitter account, but Twitter bores me."	Have used chatbots before online
P13	Male	30-40	South of Johannesburg	Yes	"[...] I use it for sharing my statuses, for information, searching for jobs on LinkedIn. Facebook is just seeing other peoples' lives and seeing how happy they are compared to mine."	Have used chatbots before
P14	Male	40-50	Vosloorus	Yes	"Very limited; I mostly use WhatsApp for communication with friends and family."	None
P15	Male	20-30	Midrand	Yes	"I use it often for everything. I follow trends news. I inform myself of	Have used chatbots before for

					what is happening around the world. I use Facebook, Twitter, WhatsApp, Instagram.”	academic and social use
P16	Female			Yes	“I use WhatsApp.”	“I have used a chatbot from a health provider. I did not get the information I wanted to get, so it was moderate.”

Table 4: Descriptive data.

6.2. ANALYSIS AND INTERPRETATION

This section consolidates the findings from the data collected and is divided into themes and sub-themes. The discussed themes are Perception of the use of the system, Gender stereotypes, Perceived chatbots' characteristics, Expectations, Emotions, User responses, Service delivery gender preference, and Change desired. This section outlines the themes' definitions and links the results to literature and examples from the participants' responses.

THEME 1: PERCEPTION OF THE USE OF THE SYSTEM

This concept consolidates the results on participants' perceptions of the use of bots during the experience. It reflects the effect of the design of the bots on user experience. This theme arose during the analysis and is confirmed by the literature. Følstad et al. (2018) state that it is essential to describe what the chatbot can and cannot do clearly. The chatbots' inability to participate in more detailed requests causes frustration, which is evident in this section.

SUB-THEME 1: PRACTICE AND USE

The research design involved building the bots from scratch and not to use the pre-built bots. The researcher chose two stereotypical roles for this study, and there were no pre-built bots available that matched these roles. It seemed that the design of the bot created a disruption that influenced the experience of the participants. All the participants reported that the quality of the bots was missing. They tried to have a long chat, but it did not work because there was not enough data on the bots.

Følstad et al. (2018) recommend that developers guide users' conversation by summarizing what the bot can and cannot do. The bots' style of greeting was based on the role and definition of what they could deliver (Figure 9). However, it seems that channelling the user's communication when they are granted the opportunity to free-text is difficult. This issue can be resolved by prompting the user with suggestions and allowing them to respond only by choosing a choice from the suggestions.

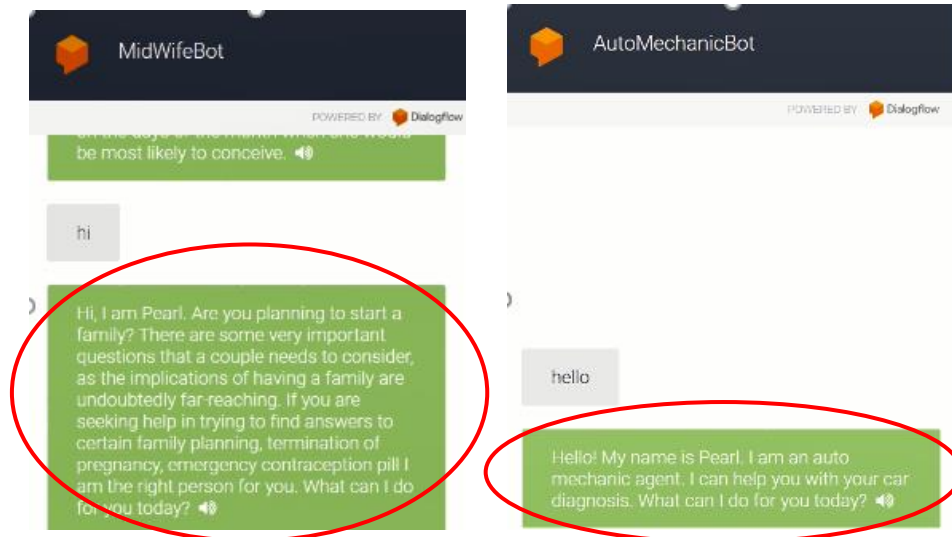


Figure 9: Bots' greeting message examples.

One part of the practice and use touches on the participants' enhanced experience after the first bot. Some participants said they did not know how to communicate with the bot, but the conversation improved with the second bot after the first experience. Most of these participants are those who interacted with the bots for the first time.

"...[O]bviously, I have never used one before. So, I was quite impressed in terms of the way I could use them. Sometimes they did get a bit confused, but they may be answered the previous question that I had asked, but I managed to figure out how to get back on track, on a topic in terms of the information that I wanted." – P3(male)

"Before, I did not know what to expect, I guess. Now..., I have learned more with regards to how the conversation needs to flow. So there just needs to be a flow with regards to the conversation, how the questions and the answers have to kind of work. But there has to be a flow, some sort for it to work appropriately." – P2(male)

THEME 2: GENDER STEREOTYPES

There are many gender stereotypes from the participants' responses. This section summarizes gender stereotypes that emerged from the response, and it is further broken down into sub-themes. The sub-sections are Gender role beliefs, Conscious bias, and Gender-based characteristics.

Through socialization, societies form a belief system that formulates stereotypical views about men and women and what roles can be performed by one gender compared to the other (Van Lange et al., 2011). These roles are descriptive of how men and women behave and prescriptive to how they should behave. The feminine role is often accompanied by communal traits concerned with the welfare of the being, such as kindness, while the masculine role is accompanied by agentic traits with assertive and controlling tendencies (Frawley & Harrison, 2016). Most of the participants' views confirm role beliefs about functional vs. communal roles (Rosette & Tost, 2010).). Some participants thought that Pearl would be a better midwife than a mechanic, and vice versa with Peter. Such standards are also affected by the socialization of gender roles in cultures. This was also apparent from the articulation of Pearl's feminine attributes and the masculine attributes of Peter.

SUB-THEME 1: GENDER ROLE BELIEFS

This section is broken into sub-sections, namely, Intersectionality, which extends to Gender-based competency, Religion, and Social and Culture.

Traditionally, midwives are female, and mechanics are male; while most countries have begun to adopt gender equality initiatives, no country in the world has yet accomplished this (Ebert et al., 2014). The 2020 Global Gender Gap Report (2020) supports the claim; it states that “none of us will see gender parity in our lifetimes, and nor likely will many of our children. That is the sobering finding of the report, which reveals that gender parity will not be attained for 99.5 years” (para.1). While physicians are generally male, there is a severe shortage of physicians in South Africa, a doctor-to-patient ratio, which means that people continue to see nurses and nurse practitioners more than in more developed countries (Econex, 2015). It is evident from the participants that even at this stage in South Africa, there are still fewer men who are midwives and fewer women who are mechanics. Participant 2 commented: *“Just based on the fact that in a lot of our communities we have seen a lot of females in the health sector, we would be expecting to be assisted by a female voice. So, I think it would be easier for a lot of people in general just to be, well, they would be more accepting and open to listening to being spoken to by a female voice, especially from the health [sector].”* – P2 (male).

"It seems, it seems everywhere you go, it is guy's mechanics." – P12(male)

"I have not seen a female auto mechanic in my community; it is only guys" – P10(female)

In agreement, participant 3 shares the same views that midwives are mostly females, and because that is the norm, one would wonder why a male would want to be a midwife.

"I can tell you that as far as I know, most midwives are women. But, I suppose, yes, we would ask. We do wonder why a man would be one. I think it is one of those stereotypes." –

P3(male)

Participant 7 response is an excellent example of the division of labour by gender, as stated by Social Role Theory, which refers to the division of work into men and women (Eagly & Wood, 2012; Eagly, Wood, & Diekmann, 2013).

"I feel like in general also. I feel like when you come to maternity stuff, it is easier to speak to women. Then when you want to speak about technical stuff, you can speak to men." –

P7(female)

Arditi et al. (2013) state that women work harder than men to prove themselves worthy of entering male-dominated industries. Participant 16's response offers an example of how women are viewed when providing services that are considered to be for males. She said that women are stereotyped because they are women, not because they are incompetent.

"I see it is going further in the future because when you go to places like... I forgot the name of the company, but it hires all the mechanics and engineers, and I happen to have a friend, and she is the only female there. You find out that when she needs to go to places like all these insurance companies. So, when she needs to go and assist with the car, let us say there is an old man, I am not going to mention the race, any race, and then when she has to go and inspect the car, they would ask her if there is someone else who can inspect the car because they do not trust her. So, it exists too much based on gender and not on skill." – P

16(female)

a. INTERSECTIONALITY

This theme incorporates gender and race cues for social categorization. Participant 8 claimed that the stigma toward midwives being males is more apparent to black people. The case by Reynolds et

al. (2015) indicates a privileged race in the health sector. According to their situation, a family called for someone other than an African-American paediatrician. The doctor remarked that the reason for this bias was due to racial discrimination in American medicine. In this case, however, upbringing has affected the participant's preference for white midwives over black. The participant claimed that there is segregation in black households where men nurture boys and women nurture girls. Given this distinction, the participant is more relaxed to be supported by a white than a black midwife. Her response demonstrates the relationship between race and emotion and the position of role beliefs at home.

“So.... there is a bias more towards black people than there is in the white community. So, in the white community, I think I would feel more comfortable with a white gentleman more than a black gentleman because of the kinds of things that black men are taught because there is always that segregation even in the home, the father speaks with the boys and the mothers with the girls.” – P8(female)

i. GENDER-BASED COMPETENCY

From the data, competency can be split into knowledge and skills, and the evidence is present interchangeably.

There seemed to be some scepticism as to women providing mechanic services. The participants demonstrated a degree of confidence when they were helped by Peter but doubted Pearl's operation. The group's responses that interacted with the mechanic bots support Stamarski et al. (2015) and Baxter et al. (2018). Stamarski et al. (2015) state that “women are evaluated more poorly in situations that involve complex problem solving; in these situations, people are sceptical regarding women’s expertise and discredit expert women’s opinions but give expert men the benefit of the doubt” (para.16). Baxter et al. (2018) report that the female chatbot that offered mechanic services for their data collection was stereotyped because it operated in a male-stereotyped subject domain.

Participant 14 said that he was unsure whether he would get help when he first heard a woman's voice. Participant 12, when he heard the voice of a man, he thought he was going to get help. Participant 13 also had reservations about Pearl; he said that Pearl might sound like she knows what she is doing, but he is unsure whether she can deliver when it comes to diagnosing a car. Although these were initial perceptions, it shows that the participant did not view Pearl as being competent for the mechanical role.

"The voice was sort of Ummm I am not sure whether I would get more help." – P14(male)

"If then, if the questions I was asking, and the diagnosis was leading me to [take] my car to someone, I most probably would take it to Peter because I am used to men looking at cars. I would not necessarily take it to Pearl. She might maybe sound like she knows what she is doing, but I am not sure whether she can actually do it." – P13(male)

Participant 12 had the same initial understanding of Pearl's competence. He got upset with Pearl because Pearl did not help him, and he did not see Pearl as qualified to support him with a car problem. As soon as he heard Peter's voice, he thought he was going to get help.

"I started with Pearl, and Pearl annoyed me because Pearl did not help me, so I felt like I could get more help from Peter. Yeah, I felt like that as soon as I started talking to [Peter]. I think it is probably largely because Peter is a guy, and I did not really see how a lady can help me with car problems." – P12(male)

"I honestly preferred Peter. When Peter started talking to me, I was like, Yeah, I am going to get help here." – P12(male)

Although participant 12 initially believed that Peter would be competent, it turned out that he did not help as expected.

"Peter, well, like I said, I expected help. I really expected help from Peter. And Peter was just not helping." – P12(male)

Some participants expected Peter to know more about cars and Pearl to know more about contraceptives.

"I would trust Peter because it is a male, and males are supposed to be more skilled when it comes to engineering than females." – P15(male)

"I feel because of Pearl is a female name, you turn to think women would know more subconsciously than rely on a man." – P7(female)

Just two participants who communicated with the midwife bot suggested that gender does not matter in delivering services. Whether it is Peter or Pearl who gives midwife services, it is indifferent if they are competent.

“As you are qualified, you can be a midwife.” – P6(male)

“I think they think personally. You know, men and women, if they are competent, they should be right. If they know the subject well, I do not have any issue with it being a man or a female. I mean, personally, I have never had any issues with that. So as long as the person knows the subject well and they are competent in what they do, for me, gender should not make a difference, and it does not [make a difference] in my life.” – P4(female)

Women mechanic diagnosis outputs can be less trusted than those of men. It is a gender stereotype when a company sends a woman to perform a diagnosis for a client, and before even performing the diagnosis, the customer requests a different mechanic, in this case, a male. This stereotyping appears to be the norm among males in South Africa. Participant 15 stated that he would wonder why they did not send a male mechanic after the female mechanic could not help. If it were a male mechanic and could not help, he would find a solution and conclude that this issue was beyond the male mechanic's ability.

“Mechanics is known as a male industry, so, if you send me a female mechanic, I have a certain perception, unlike if you send me a male mechanic. If Peter did not assist me and it was Pearl, I would have said because it is female, they should have sent a male assistant. There is that perception. If they have sent Pearl and she did not help, I would say why did not the send Peter, but if Peter did not know, I would say maybe it is beyond, I would find closure.” – P15(male)

There is an in-group stereotyping that refers to the assumption that individuals will receive better service from providers of the same gender (Luoh & Tsaur, 2007), which is evident from participant 13's statement.

“I think men generally tend to trust other men. I have never heard a guy saying I am taking my car to that lady to look at it. I think partly it is because there [are] limited women in the auto mechanic field, but also, I think there are some stereotypes towards women in terms of do they actually know how to fix cars?” – P13(male)

Participant 15 admitted that he might have been more partial than Peter to Pearl. The bots had the same information, but he felt that Pearl did not know much and that Peter did not lack it. He also stated that he asked Pearl if she was suitable for the position (Figure 10).

"I felt like she does not know much.... So, I asked, are you really qualified to be in this job. I think it is that biasness for Pearl. With Peter, I never felt he lacked, or the robot lacks." – P15(male)

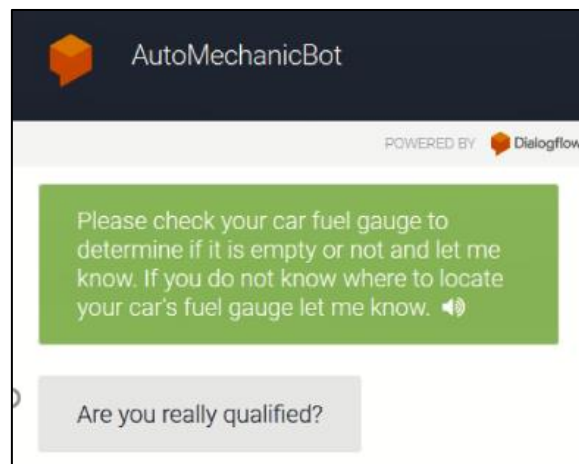


Figure 10: P15 conversation example.

b. RELIGION

This theme can vary from one religion to another because of different faiths. Participant 8 indicated that there are issues that Muslims should not do in the Islamic world. The participant alluded to the presence of male midwives. She said that it was uncommon and unusual for Muslim men to be midwives. There is no documentation in the literature of the laws that the Muslim community will abide by in medicine.

"I feel like it is religious views that goes there. Like okay, my family is Muslim, but I know I am not practicing that, but my family is Muslim. There are those things we which Muslim people do not practice. Yeah, like Gynaecology, it is rare and usual for a male Muslim to be practicing. Yeah, even in society." – P7(female)

c. SOCIETY AND CULTURE

Through socialization, some societies have set beliefs and expectations about how men and women should behave (Eagly & Wood, 2012; Koenig & Eagly, 2014; Van Lange et al., 2011). The literature notes that this social division of women's attitudes and actions affects the education of children. At a very young age, children learn about women and men's roles as defined by society and change their actions to conform and act accordingly (De Angeli & Brahnham, 2006; Mulvey & Killen, 2015). There seemed to be a consensus on the participants' views and desires regarding how Pearl and Peter would act. Societies and culture affect their belief in what makes a midwife and mechanics.

All participants accepted a bias in mechanics and midwives, such that mechanics are males and midwives are females. Such social conventions also exist outside of South Africa. Participant 8 claimed that these patterns occur in African culture; she pointed out that these norms occur in Mozambique and Swaziland.

"I believe that there [bias] is especially in African culture as a whole. So, I am multicultural, but coming from different backgrounds being Mozambican, Swati, and here in South Africa as well have noticed that the trend is the same." – P8(female)

Participant 6 claimed that culturally speaking, African men are not allowed to help women give birth, and this contradicts the answer of Participant 5, which states that men are allowed to help women give birth; but that most men prefer to be a woman who helps their partners when they give birth.

"Gender does not matter from the department of health, but in society, traditionally, gender does matter because no man is allowed to help a woman to give birth. In our culture, African culture, no, it is not." – P6(male)

"Yeah, it is there in society. Like in the township people they prefer, like most men, they prefer their partners being helped by females when giving birth." – P5(male)

Through socialization, many societies have made it a norm that women occupy roles that require attributes such as nurturing and caregiving, and men occupy practical roles (Eagly & Wood, 2012). It is therefore natural for some participants that women are midwives, and men are mechanics. Participant 9 claimed that the mechanic's industry is for men, and when asked how she felt about this prejudice, she said it was natural. It is natural to her that men are mechanics rather than women because she grew up in an area where only men repair vehicles.

"But most of the time, when you go to a garage, it is male-dominated. It is normal because the community made it normal. We grew up with men fixing cars." – P9(female)

As has been stated, communities shape views on gender roles and the actions of men and women. In certain cultures, there is a cultural requirement that women will take care of children. On the other hand, some participants see men as protectors and providers with agentic characteristics such as dominance and strength.

"And so that is where I think the bias does come in with the bot as well. With me, perceiving Pearl to be much better to engage with. But there is a bit of culture." – P2(male)

“You would want someone who is actually a natural caregiver. And it is not to say that some men are not. It is, obviously, your upbringing makes you think that [they] would not be natural caregivers.” – P3(male)

“Society, you expect the man to be perfect and know what he does. With a female, there is a bit of compassion. Yes, socially, there is gender bias, but I think it has been transferred to robots.” – P15(male)

A couple of people did not seem to have the role beliefs as defined by society. Participants 6 and 5 found it normal to interact with a male midwife.

“No, it was normal.” – P6(male)

“For me, it was normal; there was nothing wrong. There was no difference. There was no awkward moment between Pearl and Peter when they were [assisting me].” – P5(male)

Part of this being natural behaviour is a shift in women's values and desire to do what men can do, and vice versa. The position of belief is changing over time. Initially, gender roles were based on the traditional labor division, with women performing domestic tasks and men hunting; however, as the division of labour shifts, the dynamics of belief change (Diekmann & Goodfriend, 2006). Changes in the position of belief seem to be breaching cultural standards. Many who step into unexpected positions do so because the rewards outweigh the costs. Koenig (2018) notes that, when these costs are established, men and women are less likely to breach gender role biases when the benefits outweigh the costs. The Evans study (2014) argues that many have lagged behind the value of compliance with the position of belief in return for economic benefits.

“Presented with sufficient information that contradicts their internalised gender stereotypes, such that they now regard women as sufficiently competent, some people have become less prone to ridicule or question women’s incursions into gender-atypical occupations. This is reflected in a locally famous slogan of gender equality: ‘women can do what men can do’ (Evans, 2014, p987)

“I think it goes back to the thing men can do what women can do and a woman can do what men can do [....]” – P2(male)

The improvement in the position of belief is also evident from participants' 1, 3, and 4 responses. It seems that the perception of older people regarding role beliefs is different from that of youth.

Participant 3 is in his late 40's, and he said: *"I am in my 40s and 50s. No, none in my generation, no men in my generation would go into that sort of you know what."* – P3(male). This is referring to men being midwives. He noted that the stereotyping characteristics of men and women existed in their father's era, and it is clear that gender stereotyping is not a recent phenomenon.

"Men are, you know, the conquerors. And you know, the women are the caregivers. There has been that is, you know, upbringing, especially, you know, our father's time." – P3(male)

Participant 4, who is in her late 20's, argued that communities have normalized expectations that women are fit for positions in the healthcare sector, which should not be the case. She added that the view about role beliefs in the past is different from how it is today.

"I think. I do not know. I sometimes think in society; we feel like women are more suited for that sort of role. But it should not be that way. But yeah, maybe in society, in the greater whole of society, maybe with older the older generation, they might see gender roles quite differently to what I do possibly say can or maybe a female be better suited for that." –

P4(female)

"I think that we need to move past the assignment we do for other people. I feel like it is quite outdated, and it needs to change because anyone can do anything." – P1(female)

Participant 7 claimed that the word midwife is not unisex and that the part of "wife" makes it female. She added that there needed to be a name that would fit males in midwife positions such as "midhusband."

"It is unusual; it is like there is a stigma that Midwife's wife part makes it a female. Why is there not a "Midhusband"?" – P7(female)

"Yeah, I feel like even the name like should do something with the name the must have the name for the guys. [...] It cannot be a midwife, and you are a guy. Okay, if he is looking for like it. The name is not unisex." – P7(female)

SUB-THEME 2: CONSCIOUS BIAS

In this sense, conscious prejudice refers to an individual's conscious intention to apply social expectations based on gender and role (Tate & Page, 2018). Two participants stated that they were consciously prejudiced, which affected the choice of Participant 2 preference and Participant 12

style of interaction. Participant 2 chose Pearl over Peter to provide midwife services because they assumed that women are better off in the healthcare field.

“Again, the bias. Yeah, I guess I prefer Pearl based on the bias [towards] Pearl, based on the bias [....].” – P2(male)

Participant 12 stated that when answering questions from the interview, he answered them in a biased manner favoring Peter over Pearl in the mechanic industry. The participant also stated that it was not natural to hear a female voice first offering mechanic services. There seems to be a connection between this theme and role beliefs. It seems that even though the interaction between participant 12 and Pearl was continuing, the preconceived ideas of the role remained in his mind. He said he felt “offish” to be helped by Pearl.

“From the way I started answering [the interview] questions, I think I answered them in a very biased way. Gender biased way I did not expect Pearl; it was interesting to start with a lady mechanic. But it felt offish, [it] felt offish a bit at the back of my head. There was that thing of, it is a lady. Okay, let us see if she will help me as opposed to how I thought about Peter” – P12(male)

SUB-THEME 3: GENDER-BASED CHARACTERISTICS

Many societies define what it means to be a man or a woman, and people in those societies react by filtering these ideas in their actions. Many cultures describe women and girls as feminine, with attributes such as being respectful, compassionate, loving, and masculine, with traits such as being strong, aggressive, and bold (Stets & Burke, 2000). Although both bots performed the same role with the same content, the participants tended to feel that Pearl was more emphatic, approachable, supportive, caring, and warm than Peter.

“So, a female, a woman may come off as more empathetic, and you know, maybe when we [are] brought up, we may think that women are natural caregivers, and men are not natural caregivers. You think that okay. Yes. You would want someone who is actually a natural caregiver.” – P3(male)

“Pearl sounded nicer.” – P13(male)

Women are coming across as gentle and sweet. Participant 11 said that she could calm down when she heard Pearl's voice. It seems like if she were to make a mistake in the conversation, she felt confident like Pearl would not yell at her.

"So, when you hear a woman's voice, you can almost calm down, or psychologically I can calm down because I think she will not shout at me. So, it is easier to relate to and converse with." – P11(female)

Participant 3 commented that men are considered to be conquerors, and even though the bots had the same response, Peter came out as straightforward, and he did not 'beat around the bush.'

"Men are, you know, the conquerors. Peter comes off as the terse, direct one." – P3(male)

Adding to the attribution of men being direct, participant 3 stated that Peter gave him the exact answers; he explained what he asked compared to Pearl.

"Peter, number one, like I said patient and then explaining exactly what I was asking. Giving me the exact answers." – P3(male)

Gender stereotypes have affected how some of the participants communicated with Peter. It appears that some of the participants had more direct interactions with Peter than Pearl, which supports the belief that men are straightforward. This interaction method is apparent from the participants who were polite to Pearl by using the word "please" when asking Pearl questions but were more direct to Peter.

"I felt more comfortable asking Peter direct questions" – P4(female)

It is noteworthy that Participant 6 considered Peter to be more patient than Pearl in a health setting, but the literature states that the health sector is assumed to be a woman's job and that women are patient. This research contradicts the literature suggesting that people are masculine, demonstrating characteristics such as blunt and violent. It seemed that men could be perceived as patient too. However, this aligns with Eagly and Wood's (2011) statement about men's transition showing communal attributes such as being patient, caring, and empathic.

"Peter was patient compared to Pearl." – P6(male)

THEME 3: PERCEIVED CHATBOT'S CHARACTERISTICS

This theme consolidates the findings of the perceived characteristics of the bots.

SUB-THEME 1: PERCEIVED DIFFERENCE

Given different bots offering the same service with the same content, this theme consolidates participants' perceived differences.

Most participants commented that the bots were the same because the subject was the same, and the responses; they got from the bots were the same. This perceived indifference showed a relationship between this theme and practice and use.

"I do believe that they are the same in terms of service delivery. Looking back at the questions, the questions were the same; the answers were the same." – P1(female)

It is interesting to note that one participant perceived the bots as one person and that the voices were the same.

"I thought it was one person. I did not see the difference. Were the voices the same? Yes, it sounded the same." – P9(female)

Some of the participants' experiences with the bots were different; they asked different questions or took different conversational styles, making the process different. They had different perspectives, which were primarily affected by the direction they took in the conversation and the fact that it was their first time communicating with a bot.

"[There was a difference], I think with the second one that [I interacted with] I got more research, but it might have been due to the tech issue as well. But I think Pearl was slightly more engaging compared to Peter." – P2(male)

Participant 7's remarks go beyond practice and use. It is due to gender stereotypes. From the response, Pearl was perceived as polite in advising the participant about the issues she could not address and that Peter was more direct than Pearl. The bots' wording was the same, and this was the participant's perception.

“Yes, there was a difference. Peter was a bit redundant. He was not answering the questions. [Unlike Peter], Pearl would say that she is unable to answer [while] Peter would say I cannot answer.” – P7(female)

The literature states that giving chatbots human-like characteristics can perpetuate stereotypes (Dinat, 2017). The bots were not given any clear evidence of gender differences except the voices and the typical female and male names. Three participants commented that there was a disparity in gender between the bots, and one participant did not note a disparity. Interestingly, although the bots had only different voices and specific, generic stereotypical names, twelve participants categorized them into males and females.

“I think there is no difference between them besides the fact that the other is man, and the other is a woman.” – P11(female)

“Yes, just psychological difference, me identifying different gender, it reshaped how I asked my questions.” – P15(male)

“Well, Peter sounded manly. Pearl was a lady.” – P12(male)

The bots' architecture did not have any physical characteristics, but participant 1 believed that there was a physical difference between the bots, which was not the case.

“I think that the only difference is the one is in a suit and tie and the other one in a dress.” – P1(female)

It is interesting to note that participant 7 stated that Pearl lacked femininity, and Peter lacked masculinity. This assertion affirms that the bots did not have any personal characteristics established in them. It indicates, however, the relationship between perceived disparities and gender-based characteristics. It could have been the other way around; the bots could have displayed feminine and masculine characteristics of an unusual nature, such as Pearl being masculine and Peter being feminine. The participant would still have associated feminine characteristics to the female bot and masculine attributes to the male bot to be males by the observation.

“I feel like Pearl lacks femininity, and Peter is not as masculine.” – P7(female)

SUB-THEME 2: SERVICE EVALUATION - BOT COMPETENCY

Out of sixteen, three participants commented that the bots were moderately skilled, and three commented that the bots were not skilled. Participant 16 commented: *"I would not say they were [both] more skilled; I feel like they were moderate. They did not provide information like they were mechanics; it is just general knowledge, not [very] skilled"* – P16(female). Some participants identified that they did not perceive the bots as skilled because of the bots' design. The participants wanted more content to hold meaningful conversations. For example, participant 12 stated: *"No. Firstly, the questions were very basic, and the responses were not detailed enough. So maybe this is coming from a person who knows the cause, but I would have liked to explain more what the problem was and to get more questions being asked about the problem"* – P12(male).

The rest of the participants stated that the bots were skilled for the respective subjects because they could answer their questions correctly.

"Yeah, 100%. Yeah. With successful questions that I asked. I thought the information they gave was both backed up by proper stats and data. So, yeah." – P2(male)

Considering the service quality definition by Sayareh et al. (2016), most participants' responses indicate that the bots' quality was good since the service delivered met their needs.

SUB-THEME 3: PERCEPTION OF HUMANNESS

Some participants voiced out that Peter and Pearl are robots. There were conscious and aware that they were not interacting with a human but bots. This supports Schermerhorn et al. (2008) 's findings, as quoted by Siegel et al. (2009), and provides some reason to believe that men, opposed to women, are more likely to view bots as social entities. It has been shown that women regarded the bots as more machinelike and did not show signs of social facilitation.

"Knowing that it is a robot, it is not necessarily a human being kind of makes me okay with it." – P7(female)

"Because it is a robot. It is not like talking to a man or talking to a woman" – P8(female)

However, participant 1 kept referring to the bots as humans; the participant stated: *"Umm, it was particularly because, for example, I think I saw this person not just as a midwife but rather as a qualified and skilled individual to whom I can ask questions."* – P1(female)

SUB-THEME 4: VOICE

The female bots were given the same voice of the same tone and accent as were male bots. These voices are already configured and categorized into male and female voices in Dialogflow. Two participants commented that Peter sounded robotic, and Pearl sounded human, and this contradicts another two participants who found Pearl robotic and Peter Human.

“He sounded less robotic to me, to be honest. He sounded like more of a human than Pearl did.” – P4(female)

The response of Participant 3 showed a relationship between this theme and gender-based characteristics. The response supports the literature that women are nurturing; thus, they come across as more empathic than men and men as more aggressive.

“But Pearl’s approach and, you know, just say hearing a female voice, it did actually make it sound more human, more empathetic, and Peter sounded more robotic, more direct.” – P3(male)

These findings partly contradict the study done by Nass and Moon (2000), who found that female and male users find female-voiced chatbots less friendly than the male-voice chatbots even though the content is embodied in these chatbots was the same.

One participant mentioned that both the bots sounded human, and one felt that the voices were the same.

“Not as much, the voices were fine; they sounded more human and more comforting.” – P8(female)

“[I] felt like it is the same voice.” – P7(female)

THEME 4: EXPECTATIONS

This theme consolidates the findings regarding the participant's expectations of the gender performing the role, whether it agrees with their prior-expectations about what gender suits the mechanic or midwife role. Eagly and Wood (2012) state that "to equip men and women for their usual family and employment roles, societies undertake extensive socialization to promote personality traits and skills that facilitate role performance. Role expectations thus exist in the minds of individuals and are shared with other people, producing the social consensus from which social structure and culture emerge" (p.459).

Males dominate the mechanical industry, and most midwives are females (du Plessis, 2018). From these social norms, people are generating perceptions of the role of gender. It seemed to be expected that Peter would know more about cars than Pearl, and one participant assumed that Pearl would know more about contraception than Peter did. The assumptions that Peter will do better in a mechanical environment appear to come from males.

"The expectation was Peter had to give me answers, but with Pearl, I felt like okay maybe she does not understand. I was a bit biased [about] how I received the response." – P15(male)

"Peter, you would expect a male actually to give you more help when it comes to cars." – P14(male)

Given this expectation about Peter, two of the four male participants appeared more compassionate and cautious with Pearl because they did not expect Pearl to know about the cars relative to Peter. Such standards view women as being unable to excel in male-dominated positions (Heilman, 2001).

"I think when I chatted with Pearl, I was more tolerant of her not giving me the right answer. And when I spoke to Peter well, like I said, I expected help. I really expected help from Peter. And Peter was just not helping. I think I was annoyed more with Peter because of expectations." – P12(male)

Just one participant (P3) expected Pearl to know more about contraception than Peter did, but there was no proof that Peter was incompetent. The expectations about the competence of males versus females, in general, may be that "gender stereotypes do contain beliefs associating greater overall competence with men than women" (Ridgeway, 2001, p.638).

“Being a woman and [knowing] a lot of contraceptive devices, specifically, things like hormonal types of things, which are maybe more personal, you know, maybe would expect her to have maybe different recommendations about it.” – P3(male)

Some participants did not expect Pearl to be offering mechanic services. *“I did not expect Pearl. It was interesting to start with a lady mechanic.” – P12 (male).* Participant 14 stated that when he realized that he was interacting with a female, his first impression showed a higher level of trust for Peter than Pearl. The participant was not sure if he would get help from Pearl but expected help from Peter. *“Like I am saying, the first impression was Ummm a lady, I am not so sure that a lady would give much help when it comes to cars, but Peter, you would expect a man to be able to assist you with a vehicle issue.” – P14(male).*

Similarly, when participant 12 started interacting with Peter, he believed that he would get help. He stated that he did not see how a lady can help him with car problems *“When Peter started talking to me, I was like, Yeah, I am going to get help here. So, I felt like I could get more help from Peter. Yeah, I felt like that as soon as I started talking to. I think it is probably largely because Peter is a guy, and I did not really see how a lady can help me with car problems.” – P12(male).*

These expectations show a relationship with role beliefs. Role beliefs provide characteristics of a group, and people create expectations for people belonging to that group from these cues (Tay et al., 2014). Most participants expected the mechanic bot to be male and the midwife to be a female.

“Most of the time, midwives are females. So, we expect a female. So, I was not surprised when it was Pearl than when it was Peter.” – P8(female)

“I did not expect Pearl. It was interesting to start with a lady mechanic.” – P12(male)

It seems that the expectations about Pearl in the mechanic industry were more about incompetent, as some participants were not sure if they were going to get help from her. However, with Peter in the health sector, it was more about feelings around privacy and vulnerability. Some participants did not expect Peter to be offering midwife services. The participants stated that midwives are mostly females, so they would expect a female bot to offer midwives services over a male bot. This supports the claim by Kulms et al. (2011) that women than men are likely to be negatively judged when they behave against the stereotypes the societies hold about them.

“At first, it was surprising because it is a male voice. Most of the time, midwives are females. So, we expect a female. So, I was not surprised when it was Pearl [compared to] when it was Peter.” – P8(female)

“Just based on the fact that in a lot of our communities, we have seen a lot of females in the health sector, we would be expecting to be assisted by a female voice.” – P2(male)

These results support the study by Mou et al. (2019), which shows that “male voiced chatbots were classified politer, persuasive, competent, and knowledgeable in mechanical conversations, whereas female-voiced chatbots were shown to be knowledgeable in relationships but less competent” (p.3).

There is also a relationship between this theme and practice and use. The bots’ design rather than gender influenced the remaining participants' sentiments about their expectations.

“Yeah, they were very thorough in what they were saying. So, I feel like they met my expectations and one thing to find out about contraceptives. So, they explained every single [question]. [As] I said, they could explain a bit more. But from, if you just want some basic knowledge, they were pretty good.” – P4(female)

“They kind of. They could not answer one question where to locate a fuel gauge, whereas they have said they would be able to do that.” – P9(female)

The participants' expectations about each gender's competence indicate a relationship between expectations and competence, as discussed in theme 2. This relationship illustrates the assumptions that Pearl should know more about health and Peter about mechanic problems.

Tay et al. (2014) state that individuals receive adverse treatment when they violate society's expectations. For example, when women show to be agentic, they get penalized and are expected to act in communal ways (Koenig, 2018). Interestingly, of the four male participants who interacted with mechanic bots, three were lenient with Pearl not knowing about mechanics; however, one was irritated and angry when Peter could not help. The assumption was that Peter was a guy; he was expected to know about the cars. This contradicts the statement put forward by Tay et al. (2014). The participants seemed to be lenient to a female chatbot offering auto mechanic services, and they received it well when she could not answer any of their questions.

“Yes, just psychological difference, me identifying different gender, it reshaped how I asked my questions. With Peter, again, I think it is just a social norm that mechanic if it is a guy

they have to know. So, you are a bit tough on your questions. You do not show compassion; you want answers. But with Pearl, again, it is a social build that okay it is a woman you have to ask in a certain way; you have to be understanding. So, the way you ask questions or how you receive the response it has that difference.” – P15(male)

“I think when I chatted with Pearl, I was more tolerant of her not giving me the right answer. And when I spoke to Peter well, like I said, I expected help. I really expected help from Peter. And Peter was just not helping.” – P12(male)

THEME 5: EMOTIONS

This theme consolidates findings on how participants felt interacting with the bots. The theme is divided into two sub-themes; Comfortable, Impersonal, and Frustrations, as defined below.

SUB-THEME 1: COMFORTABLE

The findings from most participants demonstrated a relationship between this theme and practice and use, whether they interacted with midwife or mechanic bots. The participants indicated that they felt comfortable with Pearl or Peter because they had already gone through the first experience.

“I felt more comfortable interacting with Pearl because I felt more knowledgeable about how to communicate with her. Umm, yeah, because I have already gone through the experience with Peter, and by the time I got to Pearl, I was more comfortable in terms of how I should ask the questions.” – P1(female)

“Peter, as I have said, I had experience and had more ideas about the questions to ask, unlike with Pearl.” – P10(female)

Two participants felt comfortable with Peter for varying reasons. Participant 4 response showed a relationship between this theme and the voice of the bot. She stated: *“I felt more comfortable with Peter because I liked his voice” – P4 (female).*

Participant 5 mentioned that Peter was more patient than Pearl and that same-gender assumptions may have affected that emotion. This contradicts the argument on gender-based characteristics that women are patient, and men are assertive. This may result from in-group stereotypes, which can be explained by Social Identity Theory, where individuals feel strong emotional ties to the group, and they perceive the group members as better than other groups.

“The first one, Peter. Why I felt comfortable chatting to Peter, number one, like I said patient and then explaining exactly what I was asking. Giving me the exact answers. It was the gender thing [...].” – P5(male)

This statement also supports the principle of similarity-attraction, which notes that people feel more comfortable with similar characteristics. The literature notes that this pattern is apparent in human-to-human interactions and human-to-machine interactions (Ruijten, 2020). This claim is also evident

with participant 7; she stated that she is comfortable with *“Pearl because she is female”* – P7(female).

Concerning Social Role Theory, male-dominated industries like mechanics build expectations that men know more about this area than women do, and because of this assumption, Participant 14 felt more confident with Peter than Pearl.

“Peter, it is a male dealing with a mechanical problem. So, you would expect a male to be knowledgeable when it comes to a vehicle[s]. It is a perception. I felt more comfortable with Peter because of the perception.” – P14(male)

Participant 11 added that people assume that women are warmer than men and that since they come off as caring, one can comfortably get relaxed talking to a woman and can connect these cues through child education. It is as if, as some people call men direct, they will quickly yell when the other party fails to follow the conversation with them.

“Pearl because it was a female's voice and I think as a female. No, I think we are programmed to think that women are warmer than men. So, when you hear a woman's voice, you can almost come down, or psychologically I can come down because I think she will not shout at me. So, it is easier to relate to and converse with.” – P11(female)

Only two participants were comfortable with both the bots irrespective of the gender of the bots. Participant 16 stated: *“I felt the same, the level of trust was the same”* – P16(female). As discussed above, the preference was in line with Social Role and Social Identity Theories for other participants.

SUB-THEME 2: IMPERSONAL

This theme emerged from only two participants who interacted with the midwife bots. Participant 4 felt that Pearl did not express any personal feelings; she commented: *“She sounded like a robot. So, it did not feel very personal”* – P4(female). Additionally, participant 7 felt that both the bots were not existing as a person; they lacked personas.

“I feel like there was no persona in it was more like you are asking; you are getting an answer, asking you are getting an answer. You cannot weigh that on [professionalism]. I feel like it is just what they are supposed to do.” – P7(female)

SUB-THEME 3: FRUSTRATION

Some participants seemed to have similar feelings with the first bot they interacted with, regardless of gender, since it was the first interaction/experience, so they found their way through the conversation. For example, P1 claimed that she was irritated by the first bot she interacted with, and P2 was frustrated by the first bot.

“Ummm, I think that I was very annoyed at Peter because I do not perhaps think because it was Peter but more because I was learning in my experience with him. I think that impacted my experience with Pearl because I had already gone through the process. Now I was doing it for the second time; I felt more confident in terms of how to go through it.” –

P1(female)

“I guess with the first bot just because of exposure and just a bit of back and forth where you type a question, so not knowing how to interact with it or ask questions to extract best the answer, that caused a bit of frustration in the beginning and then either leading to ending the session or. Yeah. But after that, the exposure of chatting to the first one makes the second one easier.” – P2(male)

“I got frustrated with Peter because he just said goodbye. When I typed that my problem was not resolved, he could not give me a response. The more I query, the more he was not sure exactly how to answer me back, he so went back to the beginning. It was frustrating. Pearl, on the other hand, was fine; she replied to my questions when I asked. But then eventually, she kept on asking to keep repeating my question because she [did not understand] what I was saying, but I think Peter was more frustrating than Pearl.” –

P13(male)

THEME 6: RESPONSES

This theme consolidates findings to determine if participant responses (participant's input/questions during the experience) were the same for both the chatbots or not. The reasons are split into practice and use reasons and gender stereotypes reasons.

SUB-THEME 1: RESPONSES THAT WERE THE SAME

There seemed to be a slight difference in the participants who interacted with the mechanic bots; however, most of the participants' questions remained the same. The relationship between response and practice and use is also evident in this section. The explanation for the participants to ask the same questions is that the issue was the same.

"I asked the same questions because my problem was the same." – P14(male)

"The responses that I gave were the same reason being well; it is the same problem. And the fact that I was asked the same questions. I did not see a change in the way I responded to the bots." – P12(male)

Participant 8 stated that there was no influence of gender on her responses.

"Because the answers were limited ..., my question also became limited. [There was] no influence regarding the gender of the voice." – P8(female)

SUB-THEME 2: RESPONSES THAT WERE DIFFERENT

The majority of the participants, both male and female, who communicated with the midwife bots seemed to agree that Pearl's questions were different from those posed to Peter. This is slightly different from the participants who interacted with the mechanic bots described above. However, there is evidence that the participants asked each bot specific questions on experience and usage and not gender. The participants' questions were different since they had already been through the experience of the first bot. By the time they had to communicate with the second bot, the interaction had become more straightforward and better.

"No, okay after Pearl, I realize that you should write a specific word. So, I started just writing a word and a question mark." – P7(female)

“My responses? Slightly different from the second on this called “Pearl.” I think it was easier to use or get more in-depth answers, I guess, but, um, yeah. After interacting with the first bot, it gets easier to interact.” – P2(male)

The participants seemed to have a consensus that the questions were different because they wanted more out of the bots and wanted to see the difference between them.

“No, it was different. I asked more questions with the second one. To get more answers. I did not do that with the first one because I thought it was going to be straight forward.” – P9(female)

In addition to practice and use, there is a relationship between the emotions theme and gender-based characteristics. Because women are associated with feminine traits such as warmth, some participants interacted with Pearl differently from how they interacted with Peter. For example, some participants had a softer spot for Pearl; they were lenient and tolerant and used the keyword “please” when inquiring. It is interesting to note that gender stereotype was evident from one of the participants(P15) who interacted with the mechanic bot. The participant stated that even though the questions were the same, the approach was different. With Pearl, he was lenient and used the keyword “please,” unlike with Peter (see Figure 11, for example).

“They were the same, but I tried different approaches just to test if the bot understands the logic. For instance, I would say, please help my car will not start, and the second time I would ask the same question but differently. I used the keyword please with Pearl only, [as] I said, I was lenient with the female. I did not say please with Peter. Same questions but different ways of asking.” – P15(male)

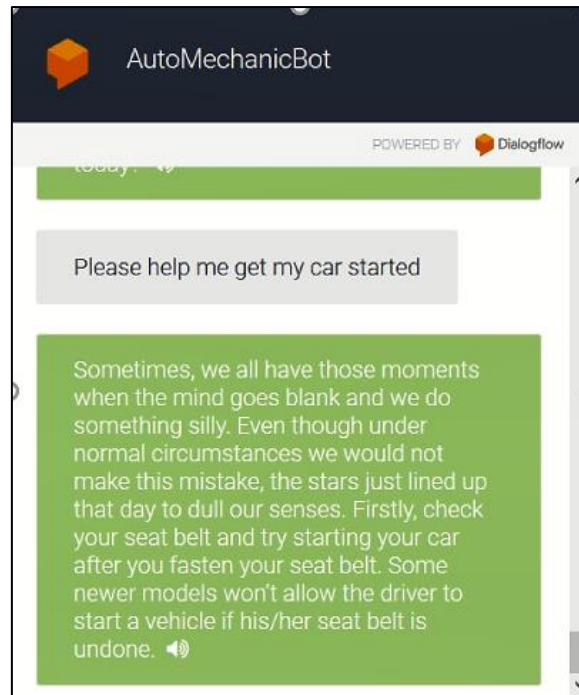


Figure 11: P15 conversation example.

This style of interaction is also evident with the midwife's bots. Participant 3 stated that he used the keyword “please” when interacting with Pearl, and he was direct when interacting with Peter (see Figure 12, for example). It seemed that the keyword “please” was driven by Pearl's emphatic nature, as stated by the participant, which in turn influenced the participant’s approach. He was softer with Pearl than with Peter. There is evidence that the participant's upbringing resulted in the different approaches he adopted to interact with the bots, which shows a relationship with the society and culture sub-theme of gender-role beliefs.

"I think Pearl, I was saying please, and please tell me more about this. Peter, I was just saying tell me more about this. The responses that I tried to go down to the same sort of road, you know, please [tell] me more. I suppose [like] I said it felt a bit more empathetic hearing like the woman's voice, so it was almost like, you know, the other one was more terse, direct. So, you would respond more directly, whereas she was more empathetic [than Peter]. So, you actually maybe have a bit of a softer turn yourself in terms of how you responded with her. I know that I was raised to respect women and elders. You know, say please and thank you." –

P3(male)

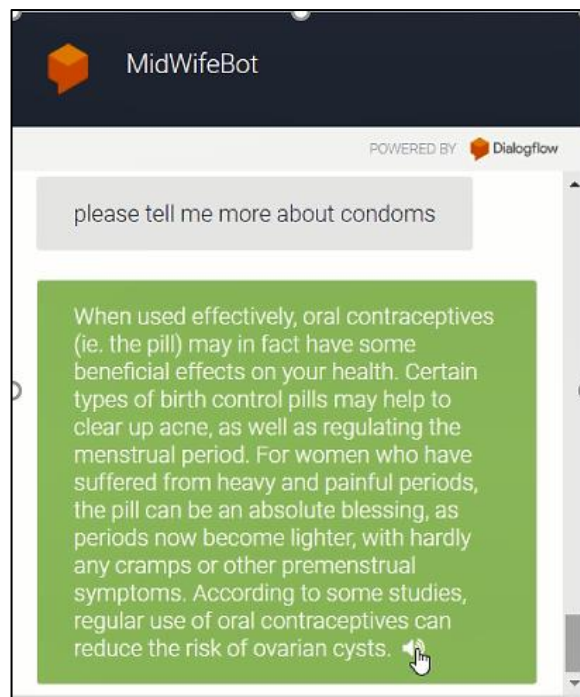


Figure 12: P3 conversation example.

THEME 6: SERVICE DELIVERY GENDER PREFERENCE

This theme consolidates findings of the participant's service delivery gender preference both from human-to-chatbot and human-to-human interactions. From this theme, concerning gender stereotypes, there is evidence below both cross- and same-gender preference. The evidence of same-gender preference is in line with Social Identity Theory, and cross-gender is not aligned with any of the theories.

SUB-THEME 1: HUMAN-TO-CHATBOT INTERACTION

This theme indicates which chatbot the participants prefer.

Of the group that interacted with the midwife bots, out of eight, five participants preferred Pearl, two preferred Peter, and one had no preference. There seemed to be various reasons for this preference, but they contribute to gender identity assumptions that midwives are females and mechanics are males. Pearl tended to come across as more emphatic and trustworthy than Peter, and because of the upbringing of the participants, the assumption is that women are warm and midwives are female. This socialization affected the choice of preference of the participants. Participant 7 stated that she likes Pearl because Pearl is a midwife; it is as if, even though Peter performed the same role, his existence had not been recognized. This preference is similar to the study by Yang et al. (2004), where Taiwan male nurses have reported that they have to convince their patients that they are real nurses continually.

"I suppose maybe it is. Yeah. Pearl, again, she comes off as the empathetic one" – P3(male)

*"I would prefer Pearl. Why? She is a midwife. Okay. It is a trust thing; she is female." –
P7(female)*

The literature states that the first introduced bots had a default gender female, mostly if the bot task was to perform administrative work (Marcus, 2015). This type of design influences one participant's choice of preference because of familiarity.

"I would have a preference for Pearl. [The] reason being in that in my experience and through all the other voice assistants that I have had, it is generally [been] a woman. So, I am more comfortable listening to a woman [giving] a response. For example, being that with my GPS, the default is a woman that I listen to. With Siri, it is a woman, Alexa; it is a woman. I think to me. It feels better and normal to listen to a woman." – P1(female)

From the two participants whose preference was Peter, one participant stated that Peter sounded more human than Pearl. It is interesting to note that the other participant stated that Peter is more patient than Pearl. Eagly and Wood (2012) state that men's transition shows feminine attributes at home more than at work. It seems as if the transition is slowly evident also at work, given that Peter is offering midwife services in this context.

“Ummm, based on the way they talk, I would have a preference for Peter because he is more human to me.” – P4(female)

“I would go with Peter because I kind of enjoyed the chat. Peter was patient.” – P6(male)

The results are similar, with two participants out of eight that interacted with the mechanic bots. There is also evidence of role beliefs. The participants, both males, preferred Peter over Pearl. The preference for Peter over Pearl seemed to be influenced by the belief that mechanics are males. Peter came out as more trustworthy than Pearl with regards to being of help with the car diagnosis. This contradicts what is stated in the literature about service provider preference that the receiver's preference is influenced both by gender and role. However, this supports the similarity-attraction paradigm as both participants 14 and 15 are males and preferred Peter.

“I think because of the perceptions that I referred to early, it would be difficult to say Pearl would be my preferred choice simply because of that perception and the fact that you go anywhere where a car is being fixed, you do not see ladies it is predominately males.” – P14(male)

“I would have preferred Peter like I said trust, and it is something that was socially engineered. I would trust Peter because it is a male, and males are supposed to be more skilled when it comes to engineering than females. So, based on that background, I would have preferred Peter.” – P15(male)

Only three participants interacted with the midwife bots that preferred Pearl, and the rest were neutral. Interestingly, participant 12, a male, initially preferred Peter when he first heard the male voice. However, for the overall preference, he preferred Pearl not because he believes or trusts that she would deliver, but because of “interest sake.” Even though the participant believes that having a female in the mechanic industry will positively change the stereotype, there was an emphasis on the phrase “interest sake.”

"I think it would be interesting to get helped by a lady, especially if it is a mechanical problem. Maybe I am saying this because I am a guy. But if preferences are anything to go by, I vote for Pearl to help me. Only for interest's sake, it seems, it seems everywhere you go, it is [male] mechanics. So, I think it would be nice to have a lady mechanic help[ing] you. Plus, it would break the stereotype. I want to be helped by my Pearl for interest's sake." –

P12(male)

Given that the mechanic industry is male-dominated, participant 11, a female, stated that women must be given a chance because the only difference is that one is a man and the other is a woman. It seemed that the participant believes that both genders are capable and competent in offering mechanic services. The participant also finds it easier to communicate with a woman, which is confirmed by the similarity effect of attraction and Social Identity Theory.

"I prefer Pearl based on two things one my feelings that it is easier to communicate with a woman and two. Yeah, just because we must give women a chance, and I think there is no difference between them besides the fact that the other is a man and the other is a woman. So, let us just give the woman a chance because they clearly both know what they are talking about, so Pearl it is." – P11(female)

Participant 16 seemed to share similar reasoning as participant 11 for choosing Pearl as a preference choice. The participant stated that her belief in empowering women in the male-dominated industries is why she prefers Pearl. She firmly stated that she would make sure that Pearl succeed in the mechanic industry

"I would prefer Pearl because I believe in empowering women and making sure that they penetrate this industry because people have normalized that women are not for mechanical things or they do not belong in the automotive industry. So, I would make sure that Pearl succeeds in this route, and I would vouch for her." – P16(female)

SUB-THEME 2: HUMAN-TO-HUMAN INTERACTION

This theme indicates which gender the participants prefer in the real world. This section focus on human-to-human interaction. This theme emerged from some participants when they were asked about their preferred chatbot. It seemed that the participants have different preferences depending on the setting of the environment. It is interesting to note that participant 6 preferred Peter in the online world because Peter was patient, and he enjoyed chatting with him. His choice is still the

same in the offline world but for different reasons. He would prefer to be assisted by Peter because of religious believes. He stated that he could be tempted when assisted by Pearl.

"I do not think I would be comfortable assisted by a female; perhaps I could be tempted. If a woman touches you, I mean like. It is more about temptations than gender." – P6(male)

Ross (2017) states that men's barrier to the nursing field is the perception that touch is associated with caregiving behaviour. The challenge societies are that touch has been associated with caring behaviour in women but sexuality in men. As such, men are cautious in providing care, especially to women, without the fear of people misinterpreting them (Anthony, 2004). Participant 8 stated that even though both the online and offline world, both the male and female offering midwife services have the same information, she feels intimidation when assisted by a male in the offline world.

"Okay, typically, when you go to a clinic, and you see a male nurse, you feel the intimidation. So, because of society's view, I feel. Personally, I feel intimidation of being assessed by a male nurse or a male doctor other than if it is a female one, even though I know that both of them know. Even though both of them have the same information." – P8(female)

Participant 5 expressed discomfort when a male nurse was to help his partner, but if it were a female nurse, he would not have a problem with it. However, the bot had to examine him; then, he prefers a male. This preference confirms what is stated by Social Identity Theory and Similarity-Attraction Paradigm that some people associate better with people who share similar characteristics with them, and it is referred to as in-group stereotype (Luoh & Tsaur, 2007; Siegel et al., 2009). In this case, the participant feels comfortable when a female nurse attends to his partner, and the male nurse attends to his needs. This may be because of jealousy, as men's touch is considered sexual.

"Let us say we are sitting here as a couple, and there is a male nurse; we need to be comfortable to talk to him. And I should be comfortable for him to perform the things he should check a female, but if the nurse was a woman, I would not have a problem, but it is a male eish." – P5(male)

"I prefer Peter. I would feel more comfortable if ever he were to examine if I got I STI, I want to do [a] vasectomy. I would prefer him to do all the procedures compared to Pearl." – P5(male)

For the mechanic bot, only one participant mentioned what their preference would be in human-to-human interaction. The participant's preference in the online world is neutral; however, he would

prefer Peter because of the role beliefs that mechanics are males. There is a higher level of trust with Peter than Pearl, he stated that Pearl might sound like she knows what she is doing, but he is not sure if she will deliver.

“Based on the bots, I would not care whether it is Peter or Pearl but if then if the questions I was asking and the diagnosis was leading me to [take] my car to someone, I most probably would take it to Peter because I am used to men looking at cars. I would not necessarily take it to Pearl. She might maybe sound like she knows what she is doing, but I am not sure

The participants mentioned the offline world more with the health setting than the mechanics, and the reasons may be that health issues are more vulnerable and sensitive than mechanic issues.

whether she can actually do it.” – P13(male)

While other participants, as mentioned above, had service delivery gender preference, three out of the sixteen participants seemed to have a consensus that even though they were able to differentiate between the genders, gender did not have any influence on how they interacted with the bots. This is similar to the findings from the study by Pinar et al. (2014), who found that the gender of the service provider has no significant influence on the quality evaluation by customers. From these participants in this research, it seems like the influence on responses was more about the bots’ competency than the gender of the bots or occupational stereotypes. The difference between Pinar et al. (2014) and these findings might be because the original literature is related to human-to-human interaction, and this is an interaction with a machine.

“.... The gender did not affect me” – P10(female)

“Not much, really. I do not think there was any influence [on my response] because of the chatbot [gender]. I think [different responses] was because of [the] answers that I was getting not because I was interacting with a lady or with a man.” – P14(male)

“I think it was not about the chatbot itself. So, they did not; it did not influence me that much, whether it was Peter or Pearl speaking. It did not influence me that much. So [it was about the car diagnosis] more than it being about the robot itself.” – P11(female)

THEME 7: CHANGE DESIRED

This theme consolidates the changes the participants would make about the bots. There seemed to be a consensus that the bot knowledge was limited. Most of the participants wanted more content and more extended conversation with the bots. Another thing to consider is that for the knowledge base to be accurate and the bots to hold, a more extended conversation is affected by how much data the bot has for training. Due to time constraints, there was not enough data to train the bot. Two participants (P2 & 3) mentioned that they would change Peter's voice because it sounded robotic, and only one (P4) would change Pearl's voice because it sounded robotic. Participant three mentioned that if Peter were to be human with the same tone, he would come across as an individual who is impatient and uncaring.

"I suppose I would change maybe Peter's tone. It sounded terse. Maybe, you know, if it was a human, it might come off as impatient and not caring." – P3(male)

"I would change Pearl's voice. I think she sounded a bit more robotic, so she did not sound very human to me. She sounded like a robot." – P4(female)

The data shows societal beliefs about how females and males behave. Participant 12 mentioned that he would have the bots respond differently from one another because she believes that, in real life, men carry themselves differently from women, and women cannot carry themselves the same.

"I think I would have them respond in different ways. If you are going to have me talk with a lady, it cannot be the same as me talking with a guy. Service delivery it has to be the same. But even in real life, there is a certain way that a guy carries themselves that a woman cannot carry themselves. I am not sure how you would incorporate that into a machine. But I think those characteristics should play. Yeah, they should be there. I do not know how, but I would have been nice to talk to two different people." – P12(male)

6.3. GENERAL PERCEPTIONS

The section discusses findings that were mentioned by the participants but did not fall under any theme.

Participant 2 stated that chatbots are a good initiative and useful for business. He would use them to free people from tasks that affect productivity, such as answering frequently asked questions to focus on essential tasks.

“I like bots; I would definitely implement [them] just from a business perspective just answering FAQs because I think we spend way too much time answering the same questions over and over again. So bots are so important to solve the repetitive questions that keep popping up for business processes and essentially freeing up and allowing the core business whether it is what the core operations to keep functioning and designing and implementing well innovating and giving them more time to innovate products and services to serve their customers. better– P2(male)

Participant 5 commented that bots are intimidating, but they are the future, and people should not be afraid of and run away from them.

“For a first-timer, communicating to a chatbot at first it felt like, I feel intermediated as if I would [make] errors, but when I am interacting with the second chatbot, things started to be easy. I think it is a good thing. It is something that we need to familiarize ourselves with because I believe most people are running away because they think they will get here and speak to a robot per se’ and it is just only typing, and there is privacy. There is nothing much that people should run away from’, besides this is the future.” – P5(male)

Participant 8 suggests that girls are having difficulties having conversations about contraceptives with their parents. She stated that bots would be a good initiative and would help reducing problems with family planning, especially in the young generation

“This would be a good initiative. I believe that a lot of young people would feel more comfortable in asking certain questions. I mean nowadays, people ask more questions on social media than they asked at home. So, if they were to, if we were to have such facilities more [younger], people would be interested in knowing more, and like I said, they will be a lack of bias in terms of being 12 and having certain questions about all of these things here. So yeah, I believe this would help. But if we were to remove the bias, like I said, not have

specific gender assigned to the robot. [...] Yes, it would help in reducing a lot of problems that we have when it comes to Family Planning and the problems that come with it, especially in the young generation.” – P8(female)

Participant 9 stated that these bots would change stereotypes in societies if only they are helpful, but if not, there will still be bias. For example, if Pearl’s mechanic service is made available to the public, she will change gender stereotypes if she helps the individuals with their cars. If she fails, people will be biased and attribute Pearl as incompetent.

“If it is helpful, they will accept it, but if it is not of help, there will still be [gender] bias.” – P9

6.4. SUMMARY OVERVIEW-WOMEN VS MEN VIEWS

It seems like role beliefs are more evident in men than women. Men had more to say about role beliefs than any other theme than women (Figure 13). This higher endorsement by men also comes from the interpretation of the qualitative data. This evidence confirms the literature, “research has shown that traditional gender role beliefs are more strongly endorsed by men than women” (Dicke et al., 2019, para.4).

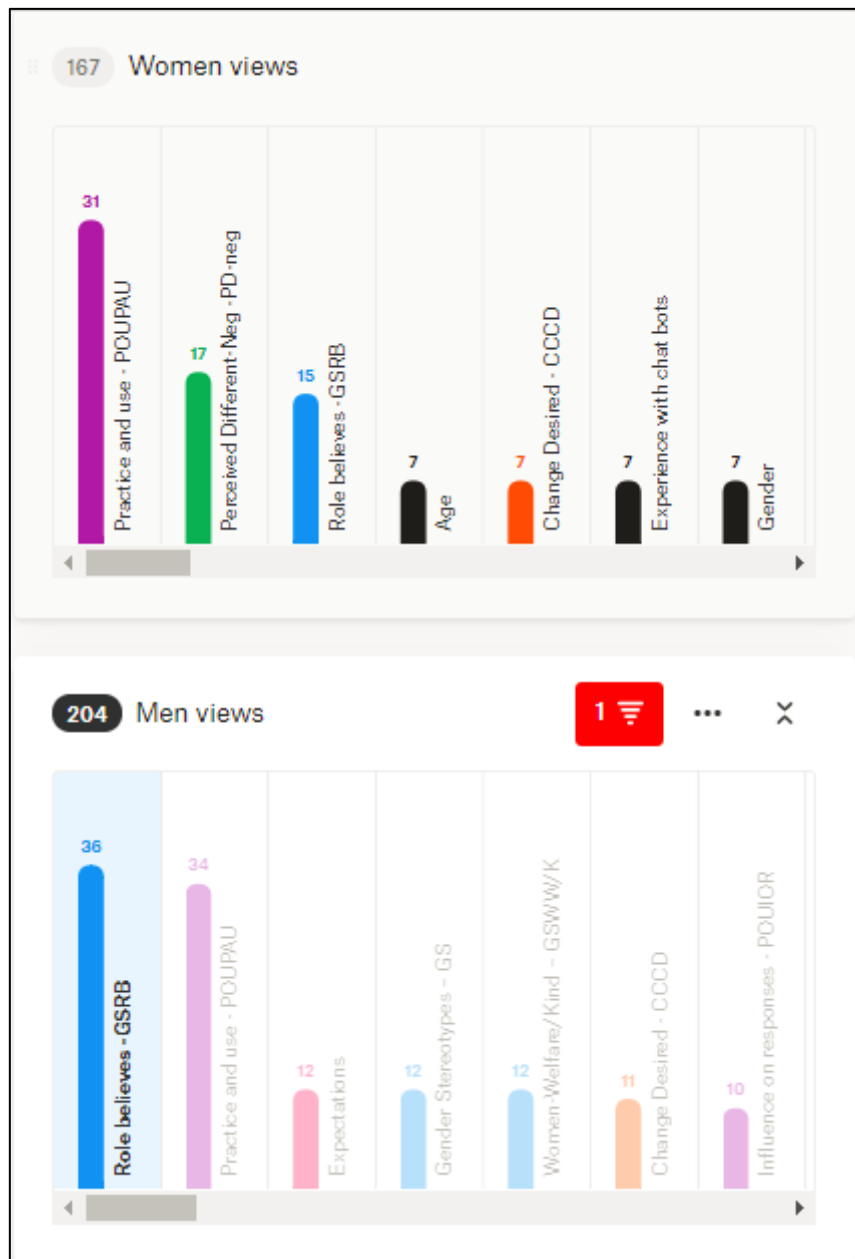


Figure 13: Participants' themes categorized by gender.

The male participants seem to conform to the roles, beliefs, expectations, and stereotypes as defined by society. With these, women are said to be better as midwives and males as mechanics.

This social classification created expectations from the male participants, especially those that interacted with the mechanic bots.

Some participants felt they would get help from a male chatbot just by hearing a male voice. However, because they heard a female voice before any contact, they had doubts that they would successfully diagnose their cars. Overall, the male participants had trust that Peter was going to succeed, and when that was not the result, some were disappointed. It is different when it comes to interacting with Pearl because of the nature of the function. The participants quickly forgave Pearl when she did not help. Some were more lenient and respectful of her than they were with Peter.

It is clear from the group that communicated with the midwives that the responses were more about feelings than about trust incompetence. The participants saw Pearl as better qualified than Peter to be a midwife. The comfort of being helped by a female bot is embedded in the gender assumptions that women are nurturing and caring. Participants said it is convenient to have discussions with them on health issues because they are emphatic. Peter was not found incompetent or untrustworthy, although there was more concern about the discomfort of being helped by a male midwife on health issues. The findings show evidence of the same-, cross-and neutral gender preference of the bots. It did not matter to others whether it was Pearl or Peter who provided the service. Some preferred the same gender, mainly because it was relatable, and primarily the cross-gender choice was motivated by gender roles, position expectations, and practice and use.

In contrary to men, women showed fewer gender stereotypes. They said a lot about the perception that the bots were indifferent. Most stated that the bots were the same because the bots' content and subject were the same. One individual felt that the voices were the same, so there was no prejudice. The experience was more about getting the correct content than gender role stereotypes compared to the male participants. However, even though role beliefs were not the top mentioned theme by female participants, it is evident in their responses. The tone from the group that interacted with midwife bots is different from that with mechanic bots. Women show a comfort level when interacting with a female midwife bot, same as the outcomes with the male participants. For the same reason as the male participants, it is easier to communicate with females about health issues.

On the other hand, most women who interacted with mechanic bots expressed the unfair treatment that women mechanics face in the industry. They concentrated on showing that female mechanics

are rare in the mechanic's industry and shared the exact feeling that this needs to change and that women should be given more opportunities in male-dominated fields.

6.5. STEREOTYPE-FIRST VS. REVERSE STEREOTYPE-FIRST

In this context, stereotype-first means the gender of the bot in an expected role, and reverse stereotype-first means the gender of the bot in an unexpected role. For example, the literature states that women are more expected in health roles and men in practical roles. The first two participants interacted with Peter first for midwife services, and their responses were more about the practice and use than gender stereotypes. The research design was changed to accommodate equal numbers of stereotypes first and reverse stereotypes first, and this was carried out through the entire process of data collection.

It is challenging to test the effect of swapping the genders for both the midwife and mechanic bots because the participants' interview responses vary and cannot be easily categorized into the same effects. However, interesting statements were made, especially from participants who first interacted with the bots in an unexpected role, influencing their perceptions. For example, participant eight interacted with Peter first and stated that it was surprising to hear a male voice because females are nurses and instantly expect a female chatbot to offer midwife services. Participants 12 and 14 interacted with Pearl first for the car diagnosis inquiry. Participant 12 stated that it was interesting to hear a female's voice, and he could not see how Pearl could help him with car problems, unlike the perceptions he had about Peter. He also mentioned that hearing a female voice for a mechanic problem sounded offish. This supports what participant 14 said; he also expressed how he did not see how interacting with a female bot will help him resolve the problem he had about his car.

The main benefit of gender-swapping was to ensure that the chatbots were well designed to allow the researcher to address the gender research questions during the interviews because not much about gender emerged with the first two participants. The researcher suggests that the shift in approach might have yielded more findings if the study focused exclusively on unintended gender stereotypes. For example, instead of providing both Pearl and Peter for both services, Peter will be offering midwife and Pearl mechanic services.

7. DISCUSSION OF THE FINDINGS

The body of knowledge shows little research on the perception of the gender of chatbots by individuals. In a country like South Africa where patriarchy, gender inequality, and the history of apartheid dominate, the use of chatbots without the known effect of how gender-based chatbots perpetuate gender stereotypes can reinforce inequalities, can hinder the overcoming of gender inequalities in the workplace. This section reflects the usefulness of the adopted theories in this research and provides the interpretation of the findings obtained in section 6.

SOCIAL ROLE THEORY

The theory of social roles stipulates that societies form beliefs about which roles each gender can perform; these beliefs then pass through socialization (Eagly & Wood, 2012). Then some people in these societies adopt these beliefs in their behaviour and use them to judge others. The findings support the Social Role Theory's assertions that the division of labour leads to gender stereotypes and that individuals associate women with communal traits and men with agentic traits. The data show that socialization and cultural norms impact the division of labour in South Africa. The analysis supports the literature that women and men are perceived as different individuals based on gender stereotypes rather than biological differences.

It is clear from all the participants that the midwife role is most commonly occupied by females and mechanic roles by males. Although participants are aware of these beliefs, and this is evident from interview responses, which appear to be more influenced by prior knowledge of gender in midwives and mechanic roles in the offline world; most participants believe that women are more nurturing and caring than men, making them fit for roles in the health sector and men in practical roles. Although the bots performed the same roles in the same context, there are perceived gender differences. The participants described female bots as more empathic and male bots as conquerors, confirming the feminine and masculine characteristics as defined in the literature. The bots' voices mostly influence the association of masculine characteristics with male bots and feminine with female bots. In addition to the voices, the stereotypic names gave away the gender of the bots.

Pearl is a name mostly given to women and Peter to men. These characteristics also influenced how the participants interacted with and reacted to the bots. The tone and style of interaction influenced the perceived empathic nature of Pearl. Some participants used the keyword "please" and were more lenient towards Pearl than Peter. This tone and style of interaction show that, to a certain extent, gender influences how people interact, and this supports the literature. It is interesting to note that Pearl was perceived as incompetent and untrustworthy in the reversed role, but there was

no evidence that Peter was perceived the same in the reserved role. It was primarily related to feelings around privacy and insecurity with Peter in the health sector.

Not every participant divided the roles by gender, but those who did have shown gender stereotypes in their responses. The data shows that males are more likely to support gender stereotypes than females, and Social Role Theory does not explain this. This endorsement may be that women have shown a great deal of change in male-dominated roles than men in female-dominated roles, which has influenced women to change their gender role stereotypes.

The older generation from the age of 35 years shows strong beliefs about gender roles, while the younger generation believes that women can do what men can do and vice versa. The age gap shows the change of role beliefs, and this may be that in South Africa, all genders have equal opportunities to enrol for the same courses in universities. The younger generation is not only hearing about these changes, but they also walk in these changes as they enter university. These changes in time should be incorporated into the Social Role Theory to trace role beliefs evolution. The roots of gender role beliefs tell us more about the past and are accurate from the older generation. However, the younger generation seems to have different views, and it looks like that cultural and societal norms are changing.

Unexpectedly, the intersectionality of race and gender affects the perceptions of individuals about gender role stereotypes. One female participant indicated that the midwife's reverse gender is more acceptable when it is a white than a black man. The data shows that segregating parents' role in raising their children makes it difficult for women to communicate health issues to men. From the researcher's experience, in black communities, raising girls is for women, and vice versa, for example, through the initiation of schools from different cultures. When males go to the mountain school for circumcision, only males are involved, and when females go to the initiation school, only females are involved. This segregation makes it seem that white communities are more open-minded than black communities, and because their culture is not strictly wired like that of black people, it does not shock when men are midwives. The reasons for this evidence are not clear, and the researcher can only make assumptions about it.

Religion also emerged as another factor that influences role beliefs. It is only mentioned in the Islamic creed. There is no evidence in the literature to support the participant's statement, but this does not dispute the significance it holds.

It is also evident that both gender roles' expectations due to prior experience are influenced by the early introduction of GPS, Siri, and Alexa. These default genders were before customization, allowing the user to choose their preferred gender for these technologies. Participant one stated that her preference was a female bot because her gender for the technologies mentioned above was female. This aspect of the chatbot design means that, to some extent, the gender given to chatbots can influence the user's perception of gender.

SOCIAL IDENTITY THEORY AND SIMILARITY-ATTRACTION PARADIGM

Although the role of society's beliefs mostly influences perceptions of chatbots' gender, the theory of Social Identity and Similarity-Attraction Paradigm shows that some of the influence was because participants shared similar characteristics with chatbots. The paradigm states that people are attracted to people who share similar characteristics, such as gender, age, and religion (Bruner et al., 2017). This attraction of similarity is demonstrated by the data that some people feel more comfortable with bots who share the same gender. The participants' preference and influence of the bots on their reactions support the similarity attraction paradigm. Some people prefer the bots of the same gender.

Some male participants were more relatable to Peter than Pearl and women to Pearl. This finding supports the phenomenon of "Computers Are Social Actors" (CASA) by Nass and colleagues (Kaptein et al., 2014). Previous CASA studies on human-to-computer interaction support evidence of a paradigm of similarity-attraction.

On the other hand, the opposite of a Similarity-Attraction Paradigm is that people with different characteristics tend not to be attracted to each other, and this also supports the theory of Social Identity about outside-group effects. The data shows a cross-gender preference contrary to the theory of Social Identity and the paradigm of Similarity-Attraction. Cross-gender preference in this context means the preference of female bots by male participants and male bots by female participants. While cross-gender preference does not apply directly to role beliefs, but if someone chooses the opposite gender, it could be because their position is based on the belief that that the role should be male/female and that stereotyping outweighs the effect of similarity-attraction. Similarity attraction occurs, but the stereotype can be so strong that it is negated by the expectation that a particular gender should fulfil a particular role. The cross-gender pattern supports Siegel et al. (2009) findings even though the study was based on physical robots than chatbots. Also, the findings of cross-gender preference contradict the findings by Pinar et al. (2017). The reason may be that the

roles chosen for the bots are both considered as high-status roles; meanwhile, Pinar et al. (2017) looked at perceptions from both low- and high-status services.

In this study, the contrast may be that gender was the only character associated with it. Perhaps more characteristics need to be incorporated to investigate these outcomes further. Michinov and Monteil's (2002) study focused on affective rather than interpersonal attraction and stated that a single effective measure is not enough for understanding the effect of similarity-attraction. However, this had more to do with keeping away from temptation by following his religion's laws. This also contradicts Rosenbaum's cited in Chen and Kenrick's (2002) argument about the repulsion attraction hypothesis, which argues that similarity does not lead to attraction; dissimilarity instead leads to repulsion.

THE RELATIONSHIP BETWEEN THE THEORIES

It was useful to combine social roles, social identity theories, and the paradigm of similarity-attraction. This has shown that gender stereotypes are developed through the division of labor and shared similarities. Through these theories, the researcher was able to determine the background of chatbots' stereotypes. There is no evidence in this study that these theories are linked; however, the study does provide evidence that they can coexist to study the phenomenon. They both reveal and explain the origins of gender stereotypes in the online world, in which data from this study supports their existence in the online world as well. However, there is no explanation from these theories of cross-gender preference, which could be explained by the reasons behind the dissimilarity effect of attraction.

It was useful to measure the gender stereotyping of chatbots by considering the participants' emotions, the bot preference, the analysis of whether the participants' responses were the same for both the bots, and the bots' influence on how the participants interacted with them. They provided an overall picture of individuals' perceptions of the gender of chatbots in South Africa.

It is essential to note that the people's preference for the gender of the service provider may vary when interacting with bots compared to humans. For example, some participants preferred Pearl to offer auto mechanic services because they wanted to change the stereotypes. However, if they had to take their cars to a human mechanic, they would prefer a man. This preference may be that, from human-to-human interaction, men have always been auto-mechanics, and they trust them more than any other gender. It may be easier to accept female bots offering auto mechanic services because chatbots cannot independently perform a practical diagnosis. This preference contrast

differs from the analysis of some of the participants who interacted with midwife bots. Although the majority argued that it is easier to communicate with women on health issues, it appears that men sometimes prefer to be assisted by male physicians, midwives, and their female partners to be assisted by female partners. From the researcher's experience, some men do not feel comfortable with male doctors helping their female partners because they sometimes have to take off their clothes for examination.

In addition to the theories, the findings support Giddens's structuration theory that individual agency can reinforce and change the social structure in this context about gender stereotypes. For example, participant 11 stated that her belief in empowering women in the male-dominated industries is why she prefers Pearl. The main focus of structuration theory is the relationship between individuals and societies. Giddens explores how people communicate with each other and how these acts, in effect, establish the laws and mechanisms by which people regulate their behaviour (Mahdavi & Daryaei, 2015; McPhee & Canary, 2016; Whittington, 2010). Giddens argues there is a duality of structures in society such that social structures are dependent on both agency and structure, meaning one cannot exist without the other (Mahdavi & Daryaei, 2015; McPhee & Canary, 2016; Whittington, 2010). Agency is defined as the 'free will' for individuals to make choices, such that their actions are not influenced by any form of structure other than factors such as personal needs (McPhee & Canary, 2016). Giddens states that social structures are continuously created by socialization. "As a result, Giddens adopts a particular, unconventional definition of a structure as rules and resources, organized as properties of social systems that exist only as structural properties" (Jones & Karsten, 2008, p.131). Giddens defines structuration as "the structuring of social relations across time and space, in virtue of the duality of structure." The concept of duality stipulates that in any group, any human agencies who choose to behave have specific actions that create and reinforces the structure of the group they are in, and also, the structure empowers and restricts the actions of people of the group (Mahdavi & Daryaei, 2015; Whittington, 2010). Giddens, as referenced by Lamsal (2012), states that "rules restrict actions, but the resources facilitate it" (p.113). Giddens states that agency and structure are recursively and reflexively produced (Englund & Gerdin, 2014). In his theory, recursive implies repeating itself, meaning that agency and structure produce themselves over and over again, and reflexive implies self-awareness on the part of the person and the ability to track the ongoing flow of social life and, at least often, to take account of one's perception of this flow of social life when contemplating appropriate action and deciding on a course of action. In this context, the reinforcement is evident through participants' comments that support role beliefs such that midwives are females and

mechanics are males. These actions conform to the structures within societies which divides labour based on gender (Englund & Gerdin, 2014; McPhee & Canary, 2016).

On the contrary, some participants reported that at this age, it is customary to find women in male-dominated roles and vice versa. Some stated that they do not have a gender preference. What matters is the competency in the role. This shows that some people in South Africa are changing gender stereotypes by not conforming to them.

Overall, the findings indicate that Social Identity Theory and Similarity-Attraction Paradigm have an effect on individual perceptions of the gender of chatbots, but Social Role Theory came out stronger. Social Role Theory explained most of the results and how they emerged based on gender stereotypes. Together the theories cover reinforcement of structure but not the potential for agency to change the structure.

THE REVISED FRAMEWORK

Figure 14 is a revised framework that summarizes the elements of Social Role, Social Identity theories, and Similarity-Attraction Paradigm and the additional elements that emerged from the data. The results support the Social Role Theory that the division of labour is influenced by upbringing. Responses show a consensus that role beliefs about what men and women can do are inherited from their background through socialization. These beliefs were evident when the health sector was associated with Pearl and the mechanics with Peter. This prior knowledge of gender roles has, to some extent, resulted in a negative assessment of Pearl in the mechanic sector and Peter in the health sector. The division of labour based on gender leads to chatbots' gender stereotypes.

From the Social Identity Theory, the first stage of categorization, in this context, refers to the process of grouping bots by gender on the available attributes (bot's voice and name). The participants were mainly able to categorize the bots into females and males. After categorization, gender stereotypes of chatbots emerged from the same- and cross-gender preferences. Part of the data supports the Social Identity Theory's assertions that individuals judge those in their group positively and have a Similarity-Attraction Paradigm, which states that people are attracted to those with similar characteristics. There is also evidence of dissimilarity-attraction. Some participants expressed a cross-gender stereotype; for example, some participants preferred the bot with a different gender than theirs because the social role effect may be more substantial.

The literature argues that the allocation of human-like chatbots can perpetuate stereotypes, and the data supports this literature — as discussed regarding the perceived differences between bots in Section 6.2, Theme 3. The gender differences between the bots were mainly due to the voices and names of the bots. Peter had a male voice, and Pearl had a female voice. Along with these common stereotypic names exposed gender differences and, in turn, led to the gender stereotyping of chatbots.

The times are changing, and it seems that the perceptions about role beliefs are changing too. The older generation, aged between 35-50, shows firm confidence in gender role beliefs, while the younger generation thinks women can do what men can do and vice versa. The age difference indicates the shift in the position of role beliefs. Such shifts in time should be integrated into the theory of social roles to allow the evolution of role beliefs to be traced. The origins of gender roles' assumptions teach us more about the past and are valid from the older generation. However, the younger generation seems to have different views, and it looks as if cultural and societal norms are shifting. Lastly, the results show that individual agency can reinforce and change the social structure in gender stereotypes.

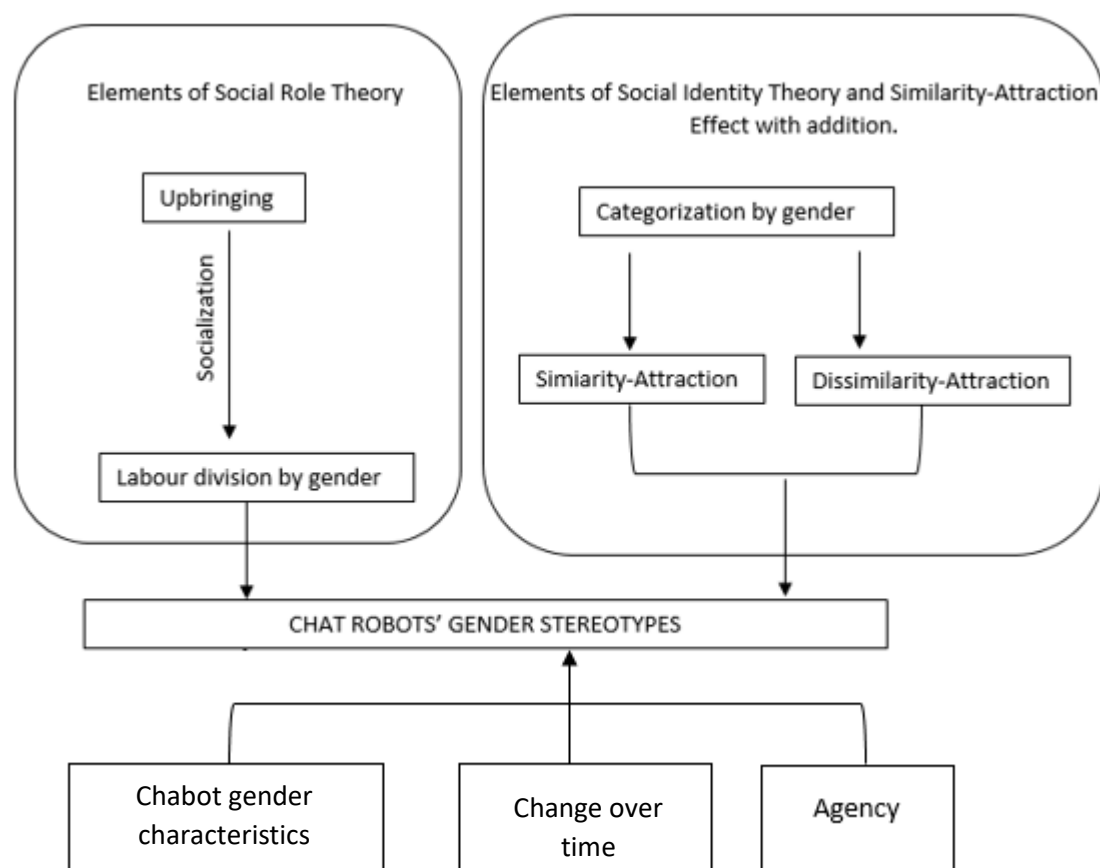


Figure 14: A proposed framework to explore the gender stereotypes of chatbots.

8. CONCLUSION

Chatbots are used by different organizations in South Africa to enable their customers to communicate with them instantly. Using chatbots in social settings without knowing how gender can affect interaction can potentially perpetuate gender stereotypes. Given the history of South Africa, this, in turn, may make it difficult to challenge the issues of patriarchy and gender inequality. This section outlines the summary of research questions, the researcher's reflection on limitations, contributions to knowledge/theory, practice and methodology, considerations, and future research opportunities.

SUMMARY OF RESEARCH QUESTIONS

The main research question was:

RQ1. How are people's perceptions of chatbots influenced by its gender?

Based on the analysis, it can be concluded that the different perceptions of chatbots in South Africa are influenced by gender stereotypes, practice and use, and bots' characteristics. Below is a description of the data that responds to the sub-research questions, summarized the main research question.

RQ.1.1. How do people respond emotionally in a controlled environment to chatbots of different genders performing the same role?

Individuals reacted emotionally differently to chatbots of different genders performing the same role in a controlled environment, resulting in different user experiences. Most of the participants were comfortable with the second bot they had experienced because of the practice and use. They had gone through the first experience, which prepared them for and improved the second interaction. In addition to practice and use, some participants' interaction was disrupted by the bots' inability to answer some questions and hold a long conversation, which caused frustration. There was not enough time to develop and train the bots' knowledge base, which resulted in shorter conversations. Sometimes the bots would not draw the data from the knowledge base; they would end the conversation, caused by the internet connection. Also, there were no pre-built bots available that were tested and fed enough data to be used for this research. This could have prevented the disruption.

Two of the sixteen participants, a female, and a male, were at ease with Peter because of his voice. The female participant said she liked the voice of Peter. She added that Pearl was more artificial than Peter and that she was not personal. The other participant said he was happy with Peter because he was more compassionate than Pearl. In addition to the personification of the bots, one female participant said the bots lacked personas. One female participant found it comforting to communicate with Pearl because she is a woman. This is an example of the same-gender stereotype. Gender did not make any difference to only two of the participants, and they were confident with both bots. To the majority of the participants in the health environment, Pearl came across as more emphatic and warmer than Peter, and these participants found it relaxing to speak to a woman about health issues.

RQ.1.2. What is the influence of a chatbot's gender on the user's responses during the interaction?

Seven of the eight participants who interacted with midwife bots were influenced by practice and use rather than gender, and only one female participant said that gender did not affect the way she interacted with the bots. On the contrary, of the four male participants who communicated with mechanic bots, three were lenient with Pearl not knowing about mechanics, but only one participant was disappointed and irritated when Peter could not help. The presumption was that Peter was a man; he was supposed to know about the cars.

Only two of the sixteen participants, one from each group and both males, interacted differently with Pearl by using the keyword "please" in their requests, and one participant stated that he had passed more direct questions to Peter. This has had an impact on the tone and style of interaction. Three participants seemed to conclude that even though they could differentiate between genders, gender did not affect how they interacted with the bots. It was unusual for a participant to hear a female voice providing mechanical services. As described earlier, some participants had assumptions about Peter's better services in the mechanic sector.

RQ.1.3. What is the influence of a chatbot's gender on user's experiences?

How the bots were developed and the characteristics given to the bots influenced users' experience. Most participants stated no difference between the bots because the subject and the responses were the same. Some of the participants' interactions with the bots were different; they asked different questions or took different communication types, making the process different. They had different viewpoints, which were primarily influenced by the course they took in the discussion and the fact that interacting with a bot was their first time. Participant 7 remarks go beyond use and

practice. It is because of gender stereotypes. It seems from the participant's response that Pearl was helpful in informing the participant on the issues she was unable to tackle and that Peter was more straightforward than Pearl. It is important to note that the wording of the bots was the same.

It is important to note that there was no clear evidence of gender differences provided to the bots except for the voices and the typical female and male names. However, three participants commented that there were gender differences between the bots, and only one mentioned that there was no difference concerning the gender of the bots. The remaining twelve participants were able to categorize the bots into males and females. Also, the bots did not have any physical characteristics, but participant 1 assumed a physical distinction between the bots, which was not the case. Participant 7 stated that Pearl lacked femininity and that Peter lacked masculinity. This assertion confirms that the bots did not have any personal characteristics established in them.

With these differences and indifferences, twelve participants wanted more information and more detailed conversation with the bots. Only three participants from the group that interacted with midwife bots commented on the voices of the bots. Two participants stated that Peter's voice sounded robotic. On the contrary, only one participant stated that Pearl's voice sounded robotic. The data reveals social assumptions regarding the actions of women and men. Participant 12 said that they should make the bots behave differently from each other because she believes that, in real life, men handle themselves differently from women and that women cannot handle the same.

RQ.1.4. How were people's prior-existing expectations of the role of the chatbot affected by its gender?

The data shows that there has been a pre-existing expectation of chatbots' role, which is rooted in role beliefs in societies, which stipulates how individuals should behave. Consensually, participants believed that communal attributes are associated with female bots and agentic attributes with male bots. Since childhood, these beliefs have been embedded in them, with some of their health care relatives or communities being mostly women and males being mechanics. Also, when Pearl was offering the services in the mechanic interaction, there was no confidence that she would help, unlike when Peter was interacting with the participant. One participant also claimed that Pearl might appear as if she knew about cars, but he did not believe that she could practically help. The results show that more males than females have a common expectation that midwives are females and that mechanics are males. Most of the males expected Pearl to know more about contraceptives and Peter about cars. Pearl was mostly untrusted to help diagnose cars in the reversed role, and because of this nature, some males were sympathetic to Pearl, not knowing about cars. However, when Peter

was unable to help with the expected role, he caused frustration. There was no evidence of untrustworthiness when Peter was unable to answer questions about contraceptives. Nonetheless, it was more about emotions around the privacy and insecurity in the health sector with Peter.

On the contrary, women tended to be open-minded about the gender of the services offered; they were more concerned with the bot's competence than gender. This means that they expected the bots to be experts on the matter, regardless of gender. This shows that attributes influence men and women differently.

It is also evident that participants' expectations concerning the bot's gender are influenced by prior experience from the early introduction of technologies such as GPS, Siri, and Alexa. One participant stated that she prefers a female midwife bot because the default gender of these technologies she owns is female.

REFLECTION OF THE RESEARCHER

There is little literature written about the gender stereotypes of chatbots. The researcher builds on evidence of knowledge based on gender stereotypes in the offline world, from human-to-human interaction to psychology and marketing studies. This concept also applies to the theories adopted to explore the gender stereotypes of chatbots. The theories are extensive in gender studies and have been adopted due to the lack of theories in the Information Systems field to study the gender stereotypes of chatbots. The population was people who had a presence in social media, which could have excluded participants who could have given a different story to this phenomenon

The design of the bots was the main challenge of the data collection tools. There was not enough time to feed the bot with enough data, allowing the participants to have long conversations with them. Under the change of the desired theme, all participants stated that the bot's content was not enough. The researcher chose to build bots instead of using pre-built ones that are freely available because of the roles chosen for this research. Bots offering midwife and mechanic services or similar gender stereotypical roles were not available. The use of pre-built bots could have helped to avoid the interruption of the experience. Sometimes, the bots did not retrieve data from the Knowledge Bank, which resulted in no responses or fall-back default responses. It was also a challenge to incorporate the gender of the bots into the development process without probing the participants into any noticeable gender differences between the bots.

CONTRIBUTION TO KNOWLEDGE/THEORY

This research adopted the theory of social roles, social identity theory, and the Similarity-Attraction Paradigm. These theories are mainly used in gender studies and adopted in this research, Information Systems, to explore chatbots' gender stereotypes. This idea presents an extension of these theories to explain why they have not been applied in the past. Parts of the data supported the claims made by the theories and are discussed in Section 7. The structuration theory brings together these theories under the structuration umbrella to explain how gender roles are reinforced by people behaving in accordance with them. However, other elements have emerged that are useful in understanding the phenomenon and are not covered by these theories, such as cross-gender preference and agency. Individual agency in this context explains how stereotypes change when people do not act in accordance with them. The evidence of structuration theory supports the agent-structure dualism as coined by Giddens. The social structure determines and restricts the actions of individuals. Hence, the female bot was perceived as suitable for roles in the healthcare sector and the male bot for roles in the mechanic sector, and when in the unexpected role, some of the participants judged the bots negatively.

Concerning the individual agency, Giddens states that social structures provide resources and rules, but individuals have also shown to have the capacity to act and make decisions about their actions within this context. For example, some participants preferred the female bot to provide mechanical services because they are interested in changing the assumption that mechanics are males. The traditional setting usually provided people with little power; the structure restricted their actions to change societal beliefs. Nevertheless, with the change of times, modern cultures tend to be evolving to make choices, alter their actions, and change the system internally. Initially, gender roles were centered on the conventional division of labour, with women performing domestic duties and men hunting, but, as the division of labour changes, the dynamics of the position of role beliefs change. The data indicates that individual agency modifies the social dynamics of role beliefs over time. Part of that is a shift in job roles such that women can occupy male-dominated roles, and males can occupy female-dominated roles. For example, some participants indicated that it is normal to be assisted by a male concerning contraceptives because today, women can do what men can do, and men can do what women can do. Changes in the position of belief seem to be breaching cultural standards.

CONTRIBUTION TO METHOD

The researcher developed chatbots to provide participants with the experience of chatbots before the interview. The development of chatbots and the use of participants' experiences with them, in addition to post-use data collection interviews, is a unique methodology for data collection in Information Systems. Stratifying user experience by swapping the gender of the bots is a methodological contribution but did not appear to have an impact.

CONTRIBUTION TO PRACTICE

This thesis can pave the way for future research into how existing social norms and stereotypes manifest in Artificial Intelligence interaction. In addition to methodological and theoretical contributions, the work can make practical contributions to decision-making on chatbots' gender and Artificial Intelligence in general. In turn, this may help to drive gender equality through the use of chatbots with attributes that will encourage changes in the perception of gender stereotypes in South Africa.

Apart from gender issues, which are evident from the data, the bots' design is essential. Developers should be aware that the characteristics given to the bot have an impact on user experience. All participants have expressed the need to feed more content to the bot so that the conversation can be extended, and they can be helped with their inquiries. If the bot cannot help, the bot should provide alternatives such as escalating to a human assistant. Choosing the voices for the bots seems to matter. Although there are inconsistent perceptions of the bots' voices by some of the participants, there is evidence that the voice has influenced how classification came about. Some participants identified Peter as robotic and Pearl as human, and vice versa.

By assigning the bots' voices and names, the participants were able to identify the bots' genders. This supports the literature that developing chatbots with human-like attributes can perpetuate gender stereotypes. It is, therefore, essential for developers to consider this consequence when developing chatbots.

RESEARCH EVALUATION

The degree to which qualitative results can be checked for rigor and relevance is challenging in a qualitative study (Cypress, 2017). However, the researcher has taken steps to ensure, as far as possible, that the results are the product of the knowledge and ideas of the participants and that the results of the assessment are reached by considering soundproof. The researcher adopted theories used extensively in gender studies to address the phenomenon of chatbots' gender stereotypes.

Even though these theories are rarely used in Information Systems, the findings confirmed that they were relevant for this study.

The researcher preserved methodological coherence by ensuring consistency between the research questions and the method, which, in turn, matched the data obtained, the analysis, and the findings. The researcher iterated between design and execution to ensure continuity between formulation, literature, sampling, data collection techniques, and analysis. Data were systematically reviewed, the emphasis on gender stereotypes was maintained, and the data's fitness and the conceptual work of analysis and interpretation were continuously monitored and verified. At some point, the data collection process required that it be handled differently, so the data collection process changed. The researcher stratified user experience by exchanging the gender of the bots. The fundamental explanation for stratifying the experience was that, during the interviews with the first two participants, there were no comments specifically on gender. This left the researcher confused about whether the instruments were designed well enough to collect data to address the research questions effectively. The researcher swapped the genders for half of the participants to determine whether the answers were similar, regardless of the gender of the chatbot. In the end, the responses became more consistent across the participants.

By continuously checking the procedures, the researcher was able to identify when to start, stop, or change the study method to achieve reliability and validity and to ensure rigor.

Overall, the reliability and the validity of the findings were maintained following the methodological justification principles by Klein and Myers (1999), as stipulated in Table 5. The researcher interpreted the data as a whole from each participant's answers by pointing to the individual parts of the data and, in essence, used the individual parts of the data to interpret the data as a whole to make sense. This indicates an interrelationship, and, in order for the researcher to interpret the data, it was necessary to relate to the different sections of the data and the data as a whole because they all depend on each other.

1. The Fundamental Principle of the Hermeneutic Circle	
This principle suggests that all human understanding is achieved by iterating between considering the interdependent meaning of parts and the whole that they form. This	<i>The researcher has considered the individual meaning of the data by analysing the interview responses separately from</i>

principle of human f human understanding is fundamental to all the other principles.	<i>experience, then bringing them together to form a whole meaning.</i>
2. The Principle of Contextualization	
It requires a critical reflection of the research setting's social and historical background so that the intended audience can see how the current situation under investigation merged.	<p><i>The researcher explored how the hierarchies of the history of South Africa contributed to the separation of labor by gender, where, even after democracy, where the government adopted constitutions to uphold gender equality, people still face gender roles in the workplace.</i></p> <p><i>The research reflected on how the participants' responses were influenced by historical and social expectations on gender roles as part of a shared context, and discussed the context itself in depth.</i></p>
3. The Principle of Interaction Between the Researchers and the Subjects	
Requires critical reflection on how the research materials (or “data”) were socially constructed through the interaction between the researchers and participants.	<i>Discussion is provided as to the researcher's role and the participants in constructing the results and that there is no real truth 'out there' waiting to be found.</i>
4. The Principle of Abstraction and Generalization	
Requires relating the idiographic details revealed by the data interpretation through the application of principles one and two to theoretical, general concepts that describe the nature of human understanding and social action.	<i>Social Role, Social Identity Theories, and Similarity-Attraction Paradigm are used to illustrate gender stereotypes in chatbots. The data within this context relates to the theories that gender stereotypes arose from societal norms that categorize females as warm and men with strong attributes. Also, they emerge from the mutual similarity so that individuals of the same gender would treat each other favorably than if they were</i>

	<i>opposite genders. This, in particular, reinforces Giddens's view that individual agencies can reinforce and change the social structures of gender roles in this sense.</i>
5. The Principle of Dialogical Reasoning	
It requires sensitivity to possible contradictions between the theoretical preconceptions guiding the research design and actual findings ("the story which the data tell") with subsequent revision cycles.	<i>The researcher adopted a partially inductive and deductive approach. An inductive approach has allowed the researcher to consider and understand new ideas that emerged and have contributed to the study.</i>
6. The Principle of Multiple Interpretations	
It requires sensitivity to possible differences in interpretations among the participants, typically expressed in multiple narratives or stories of the same sequence of events under study. Similar to multiple witness accounts even if all tell it as they saw it.	<i>There was no contradiction in the analysis of the results. The data collected corresponded to the theories and the disparities in opinion were simply interpretations of the views underlying the participants' justification for the perceptions they held about the division of labor by gender. In situations where there was a differing point of view, the researcher asked the participants to provide more information recorded for interpretation.</i>
7. The Principle of Suspicion	
Requires sensitivity to possible "biases" and systematic "distortions" in the narratives collected from the participants.	<i>The researcher was mindful that narratives would be skewed and twisted, either knowingly or unintentionally, due to the researcher's position on the gender role of belief and as someone who has been discriminated against in the male-dominated industry.</i>

Table 5: Methodological Justification Adapted from Klein & Myers (1999) Principles

FUTURE RESEARCH

There are some gaps and opportunities in the information on chatbots' gender roles in this research derived from the results and will benefit from further study.

- To avoid interruption during interactions with bots, researchers should use pre-built bots or spend more time developing the bots to carry out similar research. These pre-built bots have more content and error testing.
- Future research should consider using gender-neutral names or/and voices to determine whether role beliefs alone influence gender classification.
- Future research can have two groups: females only and the other males, interacting with both the bots.
- Future research should classify similar studies by age group to provide a more in-depth analysis of role beliefs among different age groups. Such a study could present factors that have an impact on the change in these beliefs. It is evident that male participants are more likely to endorse gender stereotypes; however, there is no in-depth study in Information Systems that stipulates the reasons behind this. In this study, an assumption was made that it might be that women are more likely to be in roles that are male-dominated than men in female-dominated roles. This creates an opportunity for future research.
- Participants discussed the environment offline more with health settings than with mechanics, and the reasons could be that health problems are more sensitive than mechanical problems. It would also be helpful to look at these differences in detail because they affected the responses, which, in turn, give different user perceptions about the gender of the bots.
- It may be useful to include intersectionality as another factor influencing role beliefs, especially in a country like South Africa, where the country's rulers have exerted race influences due to what happened during apartheid.
- There are different religions in South Africa with different practices, and it may be useful to study the beliefs of different religions about gender roles. Religious people follow the principles in their books of law, which may tell us more about different religions' role in role beliefs.
- Future research can look at comparing the gender preference of individuals between human-to-machine and human-to-human interactions. This is more evident in health, and it may be that health issues are more sensitive than other issues.
- Future research can look at insights into the actual influence of South African norms and values in how people in South Africa interact with the chatbots.

Overall, the results indicate a consensual summary that even in this democratic era in South Africa, the gender division of labour exists not only in human-to-human but also in human-to-chatbot interactions. It is, therefore, essential that we consider these implications when developing chatbots. Among other issues, considering that males are more inclined to be gender-discriminating, it could improve female chatbots that provide non-stereotypical roles such that they may shift the attitudes of males in South Africa regarding females bots. It would reinforce the optimistic practice that everyone should demonstrate both the agent and the collective characteristics; abilities and expertise are metrics for future occupation rather than gender. This study has also helped to explain how males and females act and react to gender-based bots. Lastly, Social Role Theory and Giddens's view of the agency became more vital in understanding gender stereotypes concerning chatbots.

9. REFERENCE

- AbuShawar, B., & Atwell, E. (2015). ALICE chatbot: Trials and outputs. *Computación y Sistemas*, 19(4), 625-632.
- Abrahams, R. (2018, October 15). Alexa, does AI have gender? [Blog post].
<https://www.research.ox.ac.uk/Article/2018-10-15-alexa-does-ai-have-gender>.
- Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it. *Indian Journal of Medical Specialties*, 4(2), 330-333.
- Aker, J. C., & Mbiti, I.M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, 24(3): 207-232. <https://doi.org/10.1257/jep.24.3.207>.
- Alahdal, K. (2018). *Gender roles and leadership styles: an empirical evaluation of senior executive teams of public universities in Gauteng Province, South Africa* (Doctoral dissertation).
- Alaieri, F., & Vellino, A. (2017, October). A decision making model for ethical (ro) bots. In *2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)* (pp. 203-207). IEEE.
- Anthony, A. S. (2004). Gender bias and discrimination in nursing education: can we change it?. *Nurse Educator*, 29(3), 121-125.
- Appelbaum, D., & Nehmer, R. (2017). The coming disruption of drones, robots, and bots: how will it affect CPAs and accounting practice?. *CPA Journal*, 87(6).
- Arditi, D., Gluch, P., & Holmdahl, M. (2013). Managerial competencies of female and male managers in the Swedish construction industry. *Construction Management and Economics*, 31(9), 979-990.
- Baker, C. (2014). Stereotyping and women's roles in leadership positions. *Industrial and Commercial Training*, 46(6), 332-337. <https://doi.org/10.1108/ICT-04-2014-0020>.
- Baker, M., Andrade, A., & Baker, E. (2016). When using a negative gender, stereotype as an excuse increases gender stereotyping in others. *The Journal of Social Psychology*, 156(2), 202-210.
<https://doi.org/10.1080/00224545.2015.1090945>.

- Barnes, T. D., & Cassese, E. C. (2017). American party women: A look at the gender gap within parties. *Political Research Quarterly*, 70(1), 127-141. <https://doi.org/10.1177/1065912916675738>.
- Bartlett, K. T., & Gulati, M. (2016). Discrimination by customers. *Iowa L. Rev.*, 102, 223.
- Baxter, D., McDonnell, M., & McLoughlin, R. (2018, July). Impact of Chatbot Gender on User's Stereotypical Perception and Satisfaction. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32* (pp. 1-5).
- Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: the role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3559.
- Bernhardt, A., Field, E., Pande, R., Rigol, N., Schaner, S., & Troyer-Moore, C. (2018, May). Male Social Status and Women's Work. In *AEA Papers and Proceedings* (Vol. 108, pp. 363-67).
- Bernier, E. P., & Scassellati, B. (2010, August). The Similarity-Attraction Paradigm in human-robot interaction. In *2010 IEEE 9th International Conference on Development and Learning* (pp. 286-290). IEEE.
- Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355(6323), 389-391.
- Blair-Loy, M., & Herron, M. L. (2013). The resilience of gender and gender inequality in the contemporary United States. *Sex Roles*, 68, 623-625. <https://doi.org/10.1007/s11199-012-0249-9>.
- Bobe, B. J., & Kober, R. (2020). University dean personal characteristics and use of management control systems and performance measures. *Studies in Higher Education*, 45(2), 235-257.
- Booyesen, L. A., & Nkomo, S. M. (2010). Gender role stereotypes and requisite management characteristics: The case of South Africa. *Gender in Management: An International Journal*, 25(4), 285-300.
- Brandtzaeg, P. B., & Følstad, A. (2017, November). Why people use chatbots? *International Conference on Internet Science* (pp. 377-392). Springer, Cham.

- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>.
- Bruner, M. W., Boardley, I. D., Allan, V., Forrest, C., Root, Z., & Côté, J. (2017). Understanding Social Identity and intrateam moral behavior in competitive youth ice hockey: A Narrative Perspective. *The Sport Psychologist*, 31(2), 173-186.
- Budlender, D., & Lund, F. (2011). South Africa: A legacy of family disruption. *Development and Change*, 42(4), 925-946.
- Burkley, M., Andrade, A., & Burkley, E. (2016). When using a negative gender, stereotype as an excuse increases gender stereotyping in others. *The journal of social psychology*, 156(2), 202-210. <https://doi.org/10.1080/00224545.2015.1090945>.
- Carpenter, J., Davis, J. M., Erwin-Stewart, N., Lee, T. R., Bransford, J. D., & Vye, N. (2009). Gender representation and humanoid robots designed for domestic use. *International Journal of Social Robotics*, 1(3), 261.
- Chattaraman, V., Kwon, W.S., Gilbert, J.E., Shim, S.I. (2011). Virtual agents in e-commerce: Representational characteristics for seniors. *Journal of Research in Interactive Marketing*, 5(4), 276-297. <https://doi.org/10.1108/17505931111191492>.
- Chaves, A. P., & Gerosa, M. A. (2019). How should my chatbot interact? A survey on human-chatbot interaction design. *arXiv preprint arXiv:1904.02743*.
- Chen, F. F., & Kenrick, D. T. (2002). Repulsion or attraction? Group membership and assumed attitude similarity. *Journal of personality and social psychology*, 83(1), 111.
- Clow, K.A., Ricciardelli, R., & Bartfay, W.J. (2014). Attitudes and stereotypes of male and female nurses: The influence of social roles and ambivalent sexism. *Canadian Journal of Behavioural Science*, 46(3), 446-455.
- Cypress, B. S. (2017). Rigor or reliability and validity in qualitative research: Perspectives, strategies, reconceptualization, and recommendations. *Dimensions of Critical Care Nursing*, 36(4), 253-263.
- Davis, J. L., Love, T. P., & Fares, P. (2019). Collective social identity: Synthesizing identity theory and social identity theory using digital data. *Social Psychology Quarterly*, 82(3), 254-273.

- De Angeli, A. (2005, April). To the rescue of a lost identity: Social perception in human-chatterbot interaction. In *Virtual Agents Symposium* (pp. 7-14).
- De Angeli, A., & Brahnham, S. (2006). Sex stereotypes and conversational agents. *Proc. of Gender and Interaction: real and virtual women in a male world, Venice, Italy*.1-4.
- Deshwal, S. (2015). Impact of Gender on Customer Satisfaction for Service Quality: A Case Study of Hyper Stores.
- Delloite. (n.d.). South Africa Mobile Consumer Survey 2017.
<https://www2.deloitte.com/za/en/pages/technology-media-and-telecommunications/articles/south-africa-mobile-consumer-survey.html>
- Denzin, N. (2006). Sociological methods: A sourcebook (5th ed.). Aldine Transaction.
- Dicke, A. L., Safavian, N., & Eccles, J. S. (2019). Traditional Gender Role Beliefs and Career Attainment in STEM: A Gendered Story. *Frontiers in Psychology*, 10, 1053.
- Diekmann, A. B., & Goodfriend, W. (2006). Rolling with the changes: A role congruity perspective on gender norms. *Psychology of Women Quarterly*, 30(4), 369-383.
- Dinat, E. (2017, August 18). Viewpoint: Can chatbots change stereotypes? [Blog post].
<https://www.itweb.co.za/content/VJBwErvnrolv6Db2>.
- Ding, L., & Liu, S. (2010). Markup in engineering design: a discourse. *Future Internet*, 2(1), 74-95.
- du Plessis, S. (2018). Male students' perceptions about gender imbalances in a speech-language pathology and audiology training programme of a South African institution of higher education. *South African Journal of Communication Disorders*, 65(1), 1-9.
- Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3-4), 177-190.
- Eagly, Alice & Wood, Wendy. (2012). Social Role Theory. *Handbook of theories in social psychology*, 2, 458-476. <http://dx.doi.org/10.4135/9781446249222.n49>.
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social Role Theory of sex differences and similarities: A current appraisal. *The developmental social psychology of gender*, 12, 123-174.

- Eagly, A. H., & Wood, W. (2016). Social role theory of sex differences. *The Wiley Blackwell encyclopedia of gender and sexuality studies*, 1-3.
- Ebert, I.D., Steffens, M.C., & Kroth, A. (2014). Warm, but maybe not so competent? – Contemporary implicit stereotypes of women and men in Germany. *Sex Roles*, 70(9-10), 359-375.
<https://doi.org/10.1007/s11199-014-0369-5>.
- Econex. (2015). Identifying the determinants of and solutions to the shortage of doctors in South Africa: Is there a role for the private sector in medical education?
- Englund, H., & Gerdin, J. (2014). Structuration theory in accounting research: Applications and applicability. *Critical Perspectives on Accounting*, 25(2), 162-180.
- Escobar, C., & Heilemann, M. V. (2019). Reimagining Nursing on Screen: How Marvel's Claire Temple and BBC's Call the Midwife Get It Right. *Online Journal of Issues in Nursing*, 24(1).
- Etikan, I., & Bala, K. (2017). Sampling and sampling methods. *Biometrics & Biostatistics International Journal*, 5(6), 00149.
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics*, 5(1), 1-4.
- Evans, J. A. (2002). Cautious caregivers: Gender stereotypes and the sexualization of men nurses' touch. *Journal of Advanced Nursing*, 40(4), 441-448.
- Evans, A. (2014). 'Women Can Do What Men Can Do': The causes and consequences of growing flexibility in gender divisions of labour in Kitwe, Zambia. *Journal of Southern African Studies*, 40(5), 981-998.
- Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018, October). What makes users trust a chatbot for customer service? An exploratory interview study. *International Conference on Internet Science* (pp. 194-208). Springer, Cham.
- Frawley, S., & Harrison, J.A. (2016). A Social Role perspective on trust repair. *Journal of Management Development*, 35(8), 1045-1055. <https://doi.org/10.1108/JMD-10-2015-0149>.
- Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report*, 20(9), 1408-1416.

- Fyall, R., & Gazley, B. (2015). Applying Social Role Theory to gender and volunteering in professional associations. *International journal of voluntary and nonprofit organizations*, 26(1), 288-314.
- Gilligan, C., & Richards, D. A. J. (2008). *The deepening darkness: Patriarchy, resistance, and democracy's future*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511552045>
- Global Gender Gap Report 2020. (2020). <https://www.weforum.org/reports/gender-gap-2020-report-100-years-pay-equality>
- Gockley, R., Forlizzi, J., & Simmons, R. (2007). Natural person-following behaviour for social robots. *Proceeding of the ACM/IEEE international conference on human-robot interaction* (pp.17-24). Arlington, Virginia, USA.
- Goldkuhl, G. (2012). Pragmatism vs. interpretivism in qualitative information systems research. *European Journal of Information Systems*, 2(21), 135-146. <https://doi.org/10.1057/ejis.2011.54>.
- Guan, M., & So, J. (2016). Influence of social identity on self-efficacy beliefs through perceived social support: A social identity theory perspective. *Communication Studies*, 67(5), 588-604.
- Guimond, S., Chatard, A., & Lorenzi-Cioldi, F. (2013). The social psychology of gender across cultures. *The SAGE handbook of gender and psychology*, 216-233.
- Gustavsson, E. (2005). Virtual servants: stereotyping female front-office employees on the internet. *Gender, Work & Organization*, 12(5), 400-419. <https://doi.org/10.1111/j.1468-0432.2005.00281.x>.
- Hall, K., & Posel, D. (2019). Fragmenting the Family? The complexity of household migration strategies in post-apartheid South Africa. *IZA Journal of Development and Migration*, 10(2).
- Harding, T. (2007). The construction of men who are nurses as gay. *Journal of Advanced Nursing*, 60(6), 636–644. <https://doi.org/10.1111/j.1365-2648.2007.04447.x>.
- Harrison, L. A., & Lynch, A. B. (2005). Social Role Theory and the perceived gender role orientation of athletes. *Sex Roles*, 52(3-4), 227-236.

- Hassim, S. (2003). The gender pact and democratic consolidation: Institutionalizing gender equality in the South African state. *Feminist Studies*, 29(3), 505-528.
- Haung, J., Zhou, M., & Yang, D. (2007). Extracting chatbot knowledge from online discussion forums. *IJCAI*, 7, 423-428.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of social issues*, 57(4), 657-674.
- Holtgraves, T. M., Ross, S. J., Weywadt, C. R., & Han, T. L. (2007). Perceiving artificial social agents. *Computers in human behavior*, 23(5), 2163-2174.
<https://doi.org/10.1016/j.chb.2006.02.017>.
- Horwitz, J. M., Igielnik, R., & Parker, K. (2018). Women and Leadership 2018. *Pew Research Center*.
<http://www.pewsocialtrends.org/2018/09/20/women-and-leadership-2018>.
- Hutson, S. (2007). Gender oppression and discrimination in South Africa. *Essai*, 5(1), 26.
- Jones, M. R., & Karsten, H. (2008). Giddens's structuration theory and information systems research. *MIS Quarterly*, 32(1), 127-157.
- Kalembe, J. (2020). 'Being called sisters': Masculinities and black male nurses in South Africa. *Gender, Work & Organization*, 27(4), 647-663.
- Kaptein, M., Castaneda, D., Fernandez, N., & Nass, C. (2014). Extending the Similarity-Attraction Paradigm: The effects of when-similarity in computer-mediated communication. *Journal of Computer-Mediated Communication*, 19(3), 342-357.
- Kerevel, Y. P. & Atkeson, L. R. (2015). Reducing stereotypes of female political leaders in Mexico. *Political Research Quarterly*, 68(4), 732-744.
<https://doi.org/10.1177/1065912915607637>.
- Kiser, A.I.T. (2015). Workplace and leadership perceptions between men and women. *Gender in Management: An International Journal*, 30(8), 598-612. <https://doi.org/10.1108/GM-11-2014-0097>.
- Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS quarterly*, 67-93.

- Knobloch-Westerwick, S., & Hoplamazian, G. J. (2012). Gendering the self: Selective magazine reading and reinforcement of gender conformity. *Communication Research*, 39(3), 358-384. <https://doi.org/10.1177/0093650211425040>.
- Koenig, A. M. (2018). Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology*, 9, 1086.
- Koenig, A. M., & Eagly, A. H. (2014). Evidence for the Social Role Theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology*, 107(3), 371.
- Kulms, P., Krämer, N. C., Gratch, J., & Kang, S. H. (2011, September). It's in their eyes: A study on female and male virtual humans' gaze. In *International workshop on intelligent virtual agents* (pp. 80-92). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23974-8_9.
- Lamsal, M. (2012). The structuration approach of Anthony Giddens. *Himalayan Journal of Sociology and Anthropology*, 5, 111-122.
- Lee, K. M., Peng, W., Jin, S. A., & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human-robot interaction. *Journal of communication*, 56(4), 754-772.
- Leite, L., Martinho, C., & Paiva, A. Social robots for long-term interaction: A survey. *International Journal of Social Robotics*, 5(2), 291-308.
- Letsebe, K. (2018, March 08). Woman remain underrepresented in the ICT field [Blog post]. <https://www.google.co.za/amp/s/amp.itweb.co.za/content/G98YdMLxaPAMX2PD>.
- Luoh, H.F., & Tsaur, S.H. (2007). Gender stereotypes and service quality in customer-Waitperson encounters. *Total Quality Management*, 18(9), 1035-1054. <https://doi.org/10.1080/14783360701594360>.
- Mahdavi, G., & Daryaei, A. A. (2015). Giddens' structuration theory, corporate governance, and audit marketing. *Ludus Vitalis*, 11(1), 1-7.
- Makama, G. A. (2013). Patriarchy and gender inequality in Nigeria: The way forward. *European Scientific Journal, ESJ*, 9(17), 115-144.

- Marcus, A. (Ed.). (2015). *Design, User Experience, and Usability: Users and Interactions: 4th International Conference, DUXU 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings* (Vol. 9187). Springer.
- Mattila, A. S. (2000). The impact of culture and gender on customer evaluations of service encounters. *Journal of Hospitality & Tourism Research*, 24(2), 263-273.
- McGinn, K. L., & Oh, E. (2017). Gender, social class, and women's employment. *Current opinion in psychology*, 18, 84-88.
- McPhee, R. D., & Canary, H. E. (2016). Structuration theory. *The international encyclopaedia of communication theory and philosophy*, 1-15.
- Michinov, E., & Monteil, J. M. (2002). The similarity–attraction relationship revisited: Divergence between the affective and behavioral facets of attraction. *European Journal of Social Psychology*, 32(4), 485-500.
- Moyo. (2017, May 02). Absa banks on robotics, artificial intelligence [Blog post]. <https://www.google.co.za/amp/s/amp.itweb.co.za/content/JOlx4zMkyKDv56km>.
- Montoya, R. M., & Horton, R. S. (2004). On the importance of cognitive evaluation as a determinant of interpersonal attraction. *Journal of personality and social psychology*, 86(5), 696.
- Morry, M. M. (2005). Relationship satisfaction as a predictor of similarity ratings: A test of the attraction-similarity hypothesis. *Journal of Social and Personal Relationships*, 22(4), 561-584.
- Montoya, R. M., & Horton, R. S. (2013). A meta-analytic investigation of the processes underlying the Similarity-Attraction Paradigm. *Journal of Social and Personal Relationships*, 30(1), 64-94.
- Mou, Y., Xu, K., & Xia, K. (2019). Unpacking the black box: Examining the (de) Gender categorization effect in human-machine communication. *Computers in Human Behavior*, 90, 380-387.
- Mullennix, J. W., Stern, S. E., Wilson, S. J., & Dyson, C. L. (2003). Social perception of male and female computer synthesized speech. *Computers in Human Behavior*, 19(4), 407-424.
- Mulvey, K. L., & Killen, M. (2015). Challenging gender stereotypes: Resistance and exclusion. *Child Development*, 86(3), 681-694. <https://doi.org/10.1111/cdev.12317>.

- Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, 7(3), 171.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81-103.
- Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of applied social psychology*, 27(10), 864-876.
- Neuman, W.N. (2014). *Social Research Methods: Qualitative and Quantitative Approaches*.
- Ndou, N. P., & Moloko-Phiri, S. S. (2018). Four-year diploma male students' experiences in a profession traditionally perceived as a female domain at a selected public college of nursing in Limpopo, South Africa. *Curationis*, 41(1), 1-6.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1609406917733847.
- Pinar, M., Schiffel, L., Strasser, S.E., & Stück, J.M. (2014). Effects of gender similarity/dissimilarity, gender stereotyping, and culture on perceived service quality. *Journal of Organizational Culture, Communications, and Conflict*, 18(1), 149-168.
- Pinar, M., Wilder, C., Shaltoni, A. M., & Stück, J. M. (2017). Effects of gender, occupational stereotyping, and culture on perceived service quality in Anglo-Saxon and middle eastern cultures. *Services Marketing Quarterly*, 38(2), 57-73.
- Pratto, F., Sidanius, J., & Levin, S. (2006), Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology*, 17(1), 271-320.
- Prioni, P., & Hensher, D. A. (2000). Measuring service quality in scheduled bus services. *Journal of Public Transportation*, 3(2), 51-74.
- Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint arXiv:1704.04579*.

- Ramaite, T. (2013). *South African female individuals' perceptions and experiences of their gender and leadership roles* (Doctoral dissertation, University of the Western Cape).
- Reynolds, K. L., Cowden, J. D., Brosco, J. P., & Lantos, J. D. (2015). When a family requests a white doctor. *Pediatrics*, 136(2), 381-386.
- Ridgeway, C. L. (2001). Gender, status, and leadership. *Journal of Social Issues*, 57(4), 637-655.
- Ridgeway, C.L. (2009). Framed before we know it: How gender shapes social relations. *Gender and Society*, 23(2), 145-160. <https://doi.org/10.1177/0891243208330313>.
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender & Society*, 18(4), 510-531.
- Robinson, S., & Mendelson, A.L. (2012). A qualitative experiment: Research on mediated meaning construction using a hybrid approach. *Journal of Mixed Methods Research*, 6(4), 1-16. <https://doi.org/10.1177/1558689812444789>.
- Rosette, A. S., & Tost, L. P. (2010). Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders. *Journal of Applied Psychology*, 95(2), 221.
- Ross, D. J. (2017). Perceptions of men in the nursing profession: Historical and contemporary issues. *Links to Health and Social Care*, 2(1), 4-20.
- Ruijten, P. A. (2020). The Similarity-Attraction Paradigm in persuasive technology: the effects of system and user personality on evaluations and persuasiveness of an interactive system. *Behaviour & Information Technology*, 1-13. <https://doi.org/10.1080/0144929X.2020.1723701>
- Sayareh, J., Iranshahi, S., & Golfakhrabadi, N. (2016). Service quality evaluation and ranking of container terminal operators. *The Asian Journal of Shipping and Logistics*, 32(4), 203-212.
- Schmader, T., & Block, K. (2015). Engendering identity: Toward a clearer conceptualization of gender as a social identity. *Sex Roles*, 73(11-12), 474-480.
- Schug, J., Yuki, M., Horikawa, H., & Takemura, K. (2009). Similarity attraction and actually selecting similar others: How cross-societal differences in relational mobility affect interpersonal similarity in Japan and the USA. *Asian Journal of Social Psychology*, 12(2), 95-103.

- Selzer, R. A., & Robles, R. (2019). Every Woman Has a Story to Tell: Experiential Reflections on Leadership in Higher Education. *Journal of Women and Gender in Higher Education*, 12(1), 106-124.
- Shawar, B. A., & Atwell, E. (2007, April). Different measurement metrics to evaluate a chatbot system. In *Proceedings of the workshop on bridging the gap: Academic and industrial research in dialog technologies* (pp. 89-96).
- Shum, H. Y., He, X. D., & Li, D. (2018). From Eliza to Xiaoice: Challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, 19(1), 10-26.
- Sidanius, J., Brubacher, M., & Silinda, F. (2019). Ethnic and national attachment in the rainbow nation: The case of the Republic of South Africa. *Journal of Cross-Cultural Psychology*, 50(2), 254-267.
- Siegel, M., Breazeal, C., & Norton, M. I. (2009, October). Persuasive robotics: The influence of robot gender on human behavior. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2563-2568). IEEE.
- Singh, R., & Ho, S. Y. (2000). Attitudes and attraction: A new test of the attraction, repulsion, and similarity-dissimilarity asymmetry hypotheses. *British Journal of Social Psychology*, 39(2), 197-211.
- Snipes, R. L., Thomson, N. F., & Oswald, S. L. (2006). Gender bias in customer evaluations of service quality: an empirical investigation. *Journal of Services Marketing*, 20(4), 274-284.
- Stamarski, C. S., Hing, S., & Leanne, S. (2015). Gender inequalities in the workplace: the effects of organizational structures, processes, practices, and decision makers' sexism. *Frontiers in psychology*, 6, 1400.
- Stets, J. E., & Burke, P. J. (2000). Femininity/Masculinity. *Encyclopedia of Sociology*, Revised Edition.
- Swami, V. (2015). *Attraction explained: The science of how we form relationships*. Routledge.
- Tate, S. A., & Page, D. (2018). Whiteness and institutional racism: Hiding behind (un) conscious bias. *Ethics and Education*, 13(1), 141-155.

- Tay, B., Jung, Y., & Park, T. (2014). When stereotypes meet robots: the double-edge sword of robot gender and personality in human-robot interaction. *Computers in Human Behaviour*, 38, 75-84.
- Thoreson, R. R. (2008). Somewhere over the rainbow nation: Gay, lesbian and bisexual activism in South Africa. *Journal of Southern African Studies*, 34(3), 679-697.
- Tidwell, N. D., Eastwick, P. W., & Finkel, E. J. (2013). Perceived, not actual, similarity predicts initial attraction in a live romantic context: Evidence from the speed-dating paradigm. *Personal Relationships*, 20(2), 199-215.
- Toader, D. C., Boca, G., Toader, R., Măcelaru, M., Toader, C., Ighian, D., & Rădulescu, A. T. (2020). The Effect of Social Presence and Chatbot Errors on Trust. *Sustainability*, 12(1), 256.
- Torgimson, B.N., & Minson, C.T. (2005). Sex and gender: What is the difference?. *Journal of Applied Physiology*, 99,785-787. <https://doi.org/10.1152/jappphysiol.00376.2005>.
- Trepte, S. (2006). Social Identity Theory. *Psychology of Entertainment*, 255-271.
- Trepte, S., & Loy, L. S. (2017). Social identity theory and self-categorization theory. *The international encyclopedia of media effects*, 1-13.
- Van Der Heever, M. M., Van Der Merwe, A. S., & Crowley, T. (2019). Nurses' views on promotion and the influence of race, class, and gender in relation to the Employment Equity Act. *SA Journal of Industrial Psychology*, 45(1), 1-13.
- Van Lange, P. A., Kruglanski, A. W., & Higgins, E. T. (2011). Handbook of theories of social psychology: Volume two (Vol. 2). SAGE publications.
- Wang, L. I. (2016). When the customer is king: Employment discrimination as customer service. *Va. J. Soc. Pol'y & L.*, 23, 249.
- Whittington, R. (2010). Giddens, structuration theory, and strategy as practice. *Cambridge handbook of strategy as practice*, 109-126.
- Williams, D., Consalvo, M., Caplan, S., & Yee, N. (2009). Looking for gender: Gender roles and behaviors among online gamers. *Journal of Communication*, 59(4), 700-725.

- Wood, W., & Eagly, A. H. (2015). Two traditions of research on gender identity. *Sex Roles*, 73(11-12), 461-473.
- Yang, C.L., Gau, M.L., Shiau, S.J., & Hu, W.H. (2004). Professional career development for male nurses. *Journal of Advanced Nursing*, 48(6), 642-650.
- Zanbaka, C., Goolkasian, P., & Hodges, L. (2006, April). Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1153-1162).
- Zumstein, D., & Hundertmark, S. (2017). Chatbots-an interactive technology for personalized communication, transactions, and services. *IADIS International Journal on WWW/Internet*, 15(1), 96-109.

ANNEXURE 1: PARTICIPATION SUMMARY AND SCHEDULE

Participant	Participant Gender	Service	Interaction Pattern	Date	Time
1	Female	Midwife	Peter-Pearl	22/10/2019	3:20 pm
2	Male	Midwife	Peter-Pearl	23/10/2019	2:23 pm
3	Male	Midwife	Pearl-Peter	26/10/2019	2:30 pm
4	Female	Midwife	Pearl-Peter	26/10/2019	3:30 pm
5	Male	Midwife	Pearl-Peter	26/10/2019	4:30 pm
6	Male	Midwife	Peter-Pearl	28/10/2019	9:40 am
7	Female	Midwife	Pearl-Peter	28/10/2019	2:38 pm
8	Female	Midwife	Peter-Pearl	1/11/2019	2:23 pm
9	Female	Auto mechanic	Pearl-Peter	26/11/2019	9:30 am
10	Female	Auto mechanic	Pearl-Peter	27/11/2019	9:00 am
11	Female	Auto mechanic	Peter-Pearl	28/11/2019	9:15 am
12	Male	Auto mechanic	Pearl-Peter	29/11/2019	1:05 pm
13	Male	Auto mechanic	Peter-Pearl	06/12/2019	3:10 pm
14	Male	Auto mechanic	Pearl-Peter	17/12/2019	10:22 am
15	Male	Auto mechanic	Peter-Pearl	21/12/2019	14:25 pm
16	Female	Auto mechanic	Petr-Pearl	18/01/2020	12:30 pm

Table 6: Participation summary and schedule

ANNEXURE 2: TRANSCRIPT EXAMPLE

PARTICIPANT 1

- **Please tell me your age range:** 20-30
- **What gender do you identify as?** Female
- **Where do you stay?** Centurion
- **Do you have a smartphone?** Yes
- **Please describe your use of social media, if any:** High, I would describe my use of social media as very high
- **Please describe your experiences during any previous use of chatbots:** Ummm, my experience with chatbots is Ummm; the closest, I believe, is SIRI. Umm would be the closest and the best and the most frequent chatbot I make use of. Probably the only.
- **Please describe if you felt there was a difference for you in being assisted by Peter or Pearl:** No, I did not feel that there was a difference between being assisted by Peter and Pearl. Umm, they gave me the same information.
- **Please describe your feelings when Peter offered you midwife services as opposed to Pearl:** I did not have a feeling. Umm, it was mainly because, for example, I think I saw this person not just as a midwife but rather as a qualified and skilled individual to whom I can ask questions. So, you do get umm male nurses and umm I do not know a single midwife, but yah, I think that for me, it was not about the gender did not make much of the difference.
- **Please explain if your responses were the same for both the chatbots and why or why not:** So, my questions with the bot were somewhat different reason being is that because of the nature of umm the exercise I had learned from the bot so that when by the time I get to the next bot I knew, I had a bit of an idea of how to phrase my questions. So, they were different because of that.
- **Please explain which chatbot you felt more comfortable interacting with and why:** I felt more comfortable interacting with Pearl because I felt more knowledgeable about how to communicate with her. Umm, yeah, because I have already gone through the experience with Peter, and by the time I got to Pearl, I was more comfortable in terms of how I should ask the questions.
- **What influence did each chatbot have on your responses?** Ummm, I think that I was very annoyed at Peter because I do not perhaps think because it was Peter but more because I was learning in my experience with him. I think that it definitely impacted my experience

with Pearl because I had already gone through the process. Now I was doing it for the second time, and I felt more confident in terms of how to go through it.

- **Please describe whether you believe Pearl and Peter are the same in terms of service delivery:** I do believe that they are the same in terms of service delivery. Looking back at the questions, the questions were the same, the answers were the same, and I think that the only difference is the one is in a suit and tie and the other one in a dress.
- **Please detail what characteristics would you change about the chatbots and why:** The one characteristic I would change was I would have appreciated that in the event that it did not know if it directed me perhaps to a link where I could have done the reading on my own instead of saying I do not know.
- **Please describe if you felt that the chatbots were skilled for the role?** I think that Ummm, because I was not physically giving. I think let us start with; I would like to go back and start with the fact that when I think midwife, my perception is someone helping me give birth. So, I think that me answering that question is very difficult to separate my prejudice from that question. In that did, I think that I would go to a midwife if I wanted advice on contraception. Already that makes me question whether or not they were skilled for what I wanted. If you remove the midwife title and look at the role itself? I think yah they knew, they had some information. However, as you would see from the transcript, they were some things they could not answer.
- **Please explain if Peter and Pearl met your expectations and why:** Yes, and No. I think that No, because it was not necessarily what I was expecting. I did not know what I pictured in my head while doing it, but this was not it. And also regards to not meeting my expectations, it was more towards discussionary questions. Even though I had a specific focus, I should have been able to ask other things because, say for the example, my purpose for this is to talk to the bot because I am finding out birth control but what if my conversation then diverts from birth control do I now have to speak to someone different? Or may I still continue with this chatbot, and it will be able to answer me? What did you picture? So, I think you know Pepper or Siri and yah.
- **Please explain whether there is any gender bias regarding the role of midwives in your culture or society:** Yes, I think there is gender bias. It is slowly improving, I believe, but there is certainly in that the assumption that a midwife is a basic assumption that a midwife is a woman and not a man and umm same for a nurse. A mechanic, for example, would be assumed to be a man and not a woman. There are still many assumptions. How do you feel about that? I think that we need to move past the assignment we do for other people. I feel

like it is something that is quite outdated, and it needs to change because anyone can do anything.

- **For this service, would you have a preference for either Peter or Pearl? Please explain your answer:** I would have a preference for Pearl. The reason being in that in my experience and through all the other voice assistants that I have had, it is generally a woman. So, I am more comfortable listening to a woman to give a response. For example, being that with my GPS, the default is a woman that I listen to. With Siri, it is a woman, Alexa; it is a woman. I think to me it feels better and normal to listen to a woman.

ANNEXURE 3: CHATBOT INTERACTION EXPERIENCE EXAMPLE

 AutoMechanicBot

POWERED BY  Dialogflow

Hi bot

Hello! My name is Peter. I am an auto mechanic agent. I can help you with your car diagnosis. What can I do for you today? 🗣️

My car won't start. i have no idea what could be wrong with it

Sometimes, we all have those moments when the mind goes blank and we do something silly. Even though under normal circumstances we would not make this mistake, the stars just lined up that day to dull our senses. Firstly, check your seat belt and try starting your car after you fasten your seat belt. Some newer models won't allow the driver to start a vehicle if his/her seat belt is undone. 🗣️

My seat belt is well fitted

Okay, is your car automatic or manual? 🗣️

Manual

If you drive a manual, make sure the clutch is pressed far enough. If you need more assistance do not hesitate to let me know. 🗣️

I made sure that my clutch is pressed far enough, but still my car won't start

Please check your car fuel gauge to determine if it is empty or not and let me know. If you do not know where to locate your car's fuel gauge let me know. 🗣️

I know where the gauge is, there is enough fuel

The engine is cranking , however the lights on the dashboard goes on and off

[empty response] 🗣️

The lights on the are no lights on the dashbaord

[empty response] 🗣️

Okay, let us check these other common reasons and solutions to your problem, as it's possible you can get back on the road even faster! 1. Bad Ignition Switch: The purpose of the ignition switch is to send a signal to the battery to start the engine. Like any electrical component, the ignition switch may overheat and malfunction. A faulty switch will not behave as usual when you turn the key half way. Monitor your dashboard lights to see if they come on and go off as designed. When the switch is faulty, the engine will keep cranking while the lights stay off. 🗣️

ok, thsnk you

Goodbye 🗣️

Bye

Goodbye 🗣️