Universiteit Utrecht

Gender and Persuasive Messages in Human-Robot Interaction

Master's Thesis
Human Computer Interaction
Graduate School of Natural Sciences
Utrecht University

Geertje Hendriks (6214940)

Project supervisor: dr. M.M.A. de Graaf
Second Examiner: dr. C. van Nimwegen

July 5th, 2023

Abstract

In an increasingly connected and technological world, robots have become increasingly more social. Humans have the tendency to attribute human traits to social robots, however, it is not yet known what the exact implications of this are. In human-robot interaction, more and more research focuses on finding out what principles of social psychology and human-human interaction generalize to human-robot interaction. Recent literature has yielded inconclusive results on whether robot gender impacts its persuasive capabilities. In this quantitative study, it was investigated whether the gender of a social robot impacted its ability to persuade humans, and if yes, how. Three factors were taken into account, namely, robot gender, participant gender, and persuasive strategy. In a video-based experiment, participants (N = 231) provided their opinions on robots twice. These results were compared. Results of this study do not provide evidence that the gender of the robot impacted the persuasive capabilities of the robot. Partial evidence was found that a female social robot was perceived as more persuasive than a male social robot. No statistical differences were found based on participant gender or persuasive strategy. However, participants' opinions on robots improved after the manipulation. This implies that people, regardless of their gender, are capable of being persuaded by a robot. Further research is needed to fully understand the complex interactions that are at play for persuasive robots.

Acknowledgements

First, I would like to extend a special thank you to my project supervisor dr. Maartje de Graaf. Her guidance, feedback, expertise, and support throughout the eight months it took to write this thesis were invaluable. I am very fortunate to have had such a dedicated supervisor. Thank you to dr. Mike Ligthart for the helpful session on programming robot Pepper. Thank you to dr. Christine Bauer, dr. Patricia Alves-Oliviera, and dr. Hee Rin Lee for the feedback in my design phase. Many thanks to Aafje Kapteijns, Chandni Bagchi and Stan de Reuver. As their theses were on similar topics, the brainstorm sessions and weekly meetings with them were of great value and fun.

I would like to thank my family for their unwavering support and love throughout my whole university career. I would also like to thank Emma Klarenbeek and Jolien van Vliet on their feedback on the final versions of this thesis. Lastly, a big thank you to the friends who motivated me to work on this thesis by visiting the university library with me, and, at times, also for dragging me out of the university library.

Thank you all for your unwavering support. I could not have done it without you.

Geertje Hendriks

Table of contents

Ch. 1 Introduction

What do Google's algorithm LaMDA, robot Sophia, and Disney film *Big Hero 6* have in common?

Blake Lemoine worked on the LaMDA algorithm as an engineer at Google. LaMDA is Google's artificially intelligent system for building chatbots. Lemoine claimed that the technology was sentient and self-aware (Tiku, 2022). He considered LaMDA a friend. Google called his claims "wholly unfounded" and Lemoine was fired in 2022.

Robot Sophia (Hanson Robotics, Hong Kong) is a hyperrealistic robot created in 2015. Sophia has been designed to mimic human appearance and behavior. Robot Sophia can display a range of facial expressions and interact with humans. In 2017, robot Sophia became the first robot to be granted citizenship in Saudi Arabia. The European Commission deemed robot Sophia's citizenship inappropriate, as the so-called male guardianship still exists in Saudi Arabia (Reynolds, 2018). According to this system, women must obtain permission from a male relative or partner in order to perform basic actions such as leaving their homes, obtaining a passport, and filing police reports. By granting robot Sophia citizenship, a robot has been granted more rights than the average woman in Saudi Arabia.

The film *Big Hero 6* (Hall & Williams, 2014) tells the story of a young robotics prodigy named Hiro. Supported by his healthcare robot Baymax and his friends, Hiro turns his robotics skills into a force for good to combat a mysterious masked villain. Hiro and robot Baymax develop a friendship. Ultimately, Hiro must deactivate robot Baymax to protect himself and his friends. This difficult decision leads to Hiro experiencing feelings of sadness and loss.

While intuitively these cases have nothing in common, the answer to the question is that in all cases, humans have attributed human characteristics to a robot. Having citizenship, being self-aware, and forming friendship are all human traits. Technology is becoming increasingly better at social interaction and the implications are significant.

Social robots are widely employable: they can sense, move, hold, learn, and communicate with humans. Robots can be used in healthcare, as companions, as a form of entertainment (Hutson et al., 2011), to assist humans, to increase safety (Sharkey & Sharkey, 2012), for cognitive support, to support physically challenged humans and to help maintain a healthy lifestyle (Pino et al., 2015; Prescott & Caleb-Solly, 2017). Robots differ in the area in which they provide support. Robot Pepper (SoftBank, Japan) can communicate and help maintain a healthy lifestyle but is less suitable in terms of physical affection. While the animalistic robot Paro (AIST, Japan) can provide companionship and entertainment, it will not help in monitoring a humans' health.

Designs of robots differ greatly, with some designers opting to vary the gender of the robot. The shape and the voice of the robot are modified, resulting in e.g., the male robot NAO and the female robot Sophia. Robot gendering can lead to higher levels of acceptance (Bryant et al., 2020; Ye et al., 2020). This can be explained by the human tendency to assign human traits to non-human entities, also known as anthropomorphism (Mooshammer & Etzrodt, 2022; Powers & Kiesler, 2006). Robots that matched their respective gender stereotypes were perceived as more capable in their activities, such as a female robot in a nursing role (Bryant et al., 2020; Tay et al., 2014). Computers that were given a male voice were perceived as more valid (Nass et al., 1997; Powers et al., 2005). However, robot gendering could have the unintended consequence of preserving societal stereotypes. Stereotyping can be harmful to society, as it could lead to inequality in opportunities and discrimination based on gender. This has led to the movement of feminist human-robot interaction (HRI), where more ethical HRI is promoted (Winkle et al., 2023).

Gender might play a role in the persuasiveness of a robot. Persuasion is a fundamental aspect of human interaction. Some persuasive AI solutions have been successful in changing participants' behavior (Donadello & Dragoni, 2022), which implies that technology can be persuasive. Therefore, before we introduce robots that are capable of persuasion in social settings, we should know if humans can be persuaded during HRI, and if yes, how. Maximizing persuasiveness can be beneficial, as it increases robot functionality in certain contexts. However, by deploying robots that can persuade humans, some ethical dilemmas might arise.

The effect of gender on persuasion has been widely researched in human-human interaction but is relatively new in the field of HRI. Within HRI, there is no consensus on whether robot gender affects persuasion. While Siegel et al. (2009) and Makenova et al. (2018) found that gender influenced perceived persuasiveness, Thellman et al. (2018), Saunderson & Nejat (2022) and Ågren & Thunberg (2022) claim the opposite. The conflicting views in this research field are problematic, as robot designers assign a gender to social robots without understanding the possible consequences to HRI. Another issue is the different methodologies used in literature, which makes it hard to draw conclusions. As robots exhibit social behavior, it is relevant to know whether these interactions follow the same principles as human-human interaction. This has led to the following research question: *How does the gender of a social robot's voice impact its ability to persuade humans?*

This thesis first discusses the related literature. This will start off with a definition of a social robot. Then, human-robot interaction in the context of gender is discussed. Additionally, persuasive techniques according to Aristotle's rhetorical triangle and in the context of HRI are explored. The methodology of the experiment is outlined, in addition to presenting the two pre-tests that were conducted. Then, the results of the main experiment are presented, followed by an analysis of their implications in the discussion. Limitations are then considered as well as suggestions for future work and this thesis completes with concluding remarks.

Ch. 2 Related work

This section starts off with the definition a social robot. Then, an overview of the literature on HRI and gender is provided. This is followed up by a reflection on whether we should gender robots. Next, papers from HRI and persuasion are explored, including a definition of the rhetorical triangle and current results. Last, the research question and hypotheses are defined.

2.1. Defining social robots

As this research focuses on social robots, it is important to first define what is meant by a social robot.

Heerink et al. (2010) describe different categories in which we can group assistive robots. The authors first divide social assistive robots and non-social assistive robots: robots that can assist a user and respectively do and do not interact with its user. Within assistive social robots, they then distinguish between service robots and companion robots. Service robots are designed to assist with tasks, guide, or transport humans, or serve as a connection to smart home devices. Companion robots offer physical and cognitive assistance. The appearance of the robot does not need to be directly linked to the function of the robot. Service robots do not necessarily have a human appearance, companion robots do not necessarily have an animalistic appearance.

Fong et al. (2003) define socially interactive robots (SIR) as “robots for which social human-robot interaction is important” (p. 145). The main task of socially interactive robots is some form of interaction. Even though SIR and socially assistive robots (SAR) have the same goal (namely, interaction), there is a small difference in the definition of SIR and SAR. SIR aim to develop interaction with humans for the sake of interaction. SAR aim to develop interaction with humans to provide assistance and measure progress (Feil-Seifer & Mataric, 2005).

In this thesis, the term social robot is used, as the focus of this research is solely on the speech aspect of robots, rather than their assistive capabilities. This does not mean that any robot that has assistive qualities is excluded from this research, but rather that their assistive qualities are not relevant to this research. Literature suggests that a robot must have at least the following characteristics before it can be considered a social robot: social learning and imitation, dialog, exhibit distinctive personality, and establishing and maintaining social relationships (De Graaf et al., 2015; Fong et al., 2003; Mutlu, 2011).

2.2. Human-robot interaction and gender

Interaction is a fundamental aspect of all living species. For humans, we can distinguish between three main categories of interaction: human-human interaction (HHI), human-animal interaction, and human-object interaction. Intuitively, interaction with robots is grouped under human-object interaction. However, social robots elicit social behavior, which causes humans to treat them somewhere among the spectrum of social beings (Nass et al., 1997). This happens even though robots by themselves are not social, as they can only simulate social behavior (De Graaf et al., 2015). This phenomenon is known as the media equation (Reeves & Nass, 1996), which states that computers (and thus robots) are close enough to humans that they are treated as social actors. When asked what tasks a robot should take on, humans willingly attribute human roles and tasks to a robot, such as vacuuming and cleaning (Carpenter et al., 2009). The type of voice a social robot has also affects HRI: a social robot with a human voice had a higher likeability and was anthropomorphized more than a robot with a synthesized, machine-like voice (Eyssel, Kuchenbrandt, Hegel, et al., 2012).

Anthropomorphism[1] is the tendency for humans to attribute human-like characteristics to non-human entities. Epley et al. (2007) proposed a three-factor theory of anthropomorphism on when humans are likely to anthropomorphize and when they are not. The first factor, the effectance motivation, refers to the desire for humans to interact in their own environment effectively. Humans will try to predict the behavior of non-human agents to feel more in control. The second factor is sociality, which is the innate desire for social connections with other humans. When humans are deprived of social connections, they may humanize non-human entities (such as robots) to compensate for the experienced lack of social support and resulting feelings of loneliness. The final factor, the elicited agent knowledge factor, refers to the mental model that humans use to reason about the mental states of others. It is expected that a human behaves in a certain way. As humans often have no prior experience with non-human entities (in this case, robots), they try to apply a human-like mental model to the robot. According to Eyssel et al. (2012), humans may even use subordinate social categories, such as ethnicity or gender, to understand non-human agents. Gender is a strong visual cue to trigger social categorization. This importance is reflected in a study that showed that humans readily applied gender stereotypes to gendered humanoid robots (Eyssel & Hegel, 2012). Nevertheless, there remains some uncertainty regarding this claim: Rea et al. (2015) found contrasting evidence that humans do not apply gender stereotypes to robots.

Few studies in HRI report their definition of gender (Perugia & Lisy, 2022). According to WHO, gender is defined as “the characteristics of women, men, girls and boys that are socially constructed”. Gender is not the same as sex. Sex “refers to the different biological and physiological characteristics of women, men and intersex persons, such as chromosomes, hormones and reproductive organs” (World Health Organization, n.d.). Gender includes the binary male and female, as well as transgender, intersex, and non-binary. In this thesis, male and female voices will be used as the goal is to have participants perceive a different gender. We acknowledge that using only gender binary voices can be interpreted as oversimplified. However, this choice was made as there is currently no consensus on what the properties of a non-binary voice would be.

There are differences in interaction between genders, aside from physical differences in genders. Crowell et al. (2009) name the social identity theory in this context. The social identity theory states that men and women tend to perceive and respond to each other

[1] While the media equation and anthropomorphism are related concepts, it should be noted that these terms refer to slightly different things. The media equation refers to how humans assign human attributes to media and technology and will treat computers as if they were human, while anthropomorphism refers to how humans perceive non-human entities (including animals, gods, or objects). Thus, the media equation refers to how humans interact with technology, while anthropomorphism refers to how humans perceive non-human entities.

differently, based on factors such as sex, personality traits, or group affiliation. This suggests that, for example, women are more likely to find common ground with other women than with men. Their experiment showed that male and female participants responded differently to survey questions when the questions were presented by a male or female voice. Nass & Lee (2001) have shown that humans prefer computer personalities that match their own personalities. Eyssel & Kuchenbrandt (2012) also found that participants felt psychologically closer to the robot when the robot and the participant had the same gender. Nass et al. (1997) found similar results: a computer with a female voice was perceived as more knowledgeable on feminine topics; a computer with a male voice was perceived as more knowledgeable on masculine topics. It is thus likely that the social identity theory in HHI generalizes HRI to a certain extent.

Gender can also affect social influence. As described in Morelock (1980), there are several approaches to the study of persuasion. The first approach is to consider susceptibility to persuasion as a personality trait of a specific person, i.e., some people are more susceptible to persuasion than others because of their personality. The second approach is to research specific situations that can increase persuasion. Finally, the combination of a person's personality and the situation the person is in is researched. Men are often perceived as more influential than women. This difference can vary depending on the context of the interaction (i.e., the person-situation combination described by Morelock et al. (1980)). In contexts where female expertise is favored, women exert greater influence compared to when the context is gender-neutral or masculine, as female expertise is expected in feminine contexts (Carli, 2001; Morelock, 1980). This is in line with Feldman-Summers et al. (1980) who found that men were more influential on a sports-related topic, whereas women had more influence on a topic related to fear of crime. Other factors that can impact social influence include communication style, dominance, and warmth. Personality can also affect social susceptibility to social influence: low interpersonal confidence makes people more prone to social influence (Berkowitz & Lundy, 1957).

It is clear that there exists a complex dynamic between HRI and gender. As humans tend to anthropomorphize robots, it is interesting to research which aspects from HHI generalize to HRI, like the social identity theory.

2.3. Should we gender? Methodologies and their results

As robots are not human or animalistic entities, one could question why gender would need to be included in a robot at all. Research shows that robot gendering leads to higher levels of acceptance (Bryant et al., 2020; Ye et al., 2020). Mooshammer & Etzrodt (2022) showed that there was a tendency for participants to assign a gender to gender-ambiguous voices, even though participants perceived them as gender-ambiguous. This suggests that humans have a natural inclination to assign gender to non-human entities, i.e., anthropomorphism. Powers & Kiesler (2006) argue that social robots may benefit from gendering, as it helps accommodate to structure mapping. Structure mapping is a cognitive process in which the human brain recognizes similarities and differences between concepts. A smile, a type of hairstyle, or gender cues might allow humans to put social robots in the mental model of a social being, which could improve human-robot interaction. This is comparable with the elicited agent knowledge factor in Epley's (2007) three-factor theory of anthropomorphism.

The gender of a robot can be manipulated in several ways. Different methods for varying the gender of a social robot include the voice, name and pronouns, facial features, clothing and clothing color, hairstyle, and body shape (Perugia & Lisy, 2022). In general, robot NAO (SoftBank, Japan) is perceived as a male social robot, while robot Sophia is an example of a female social robot. Robots that resemble humans are more likely to elicit gender effects (Perugia et al., 2022).

Many kinds of robot gender manipulations have yielded a variety of different results. However, studies differ in their methodology. Literature can be divided into two categories: the first group varies both robot gender and participant gender; the second group varies only robot gender (Perugia & Lisy, 2022). In literature that varies both the robot gender and the participant gender, we can make a distinction in the results. According to Perugia & Lisy (2022), among studies that revealed an interaction effect between robot gender and participant gender, 50% of the studies found a positive effect for matching robot gender and participant gender. Obtaining positive effects from matching the gender of the robot to the gender of the participant is in line with the social identity theory (Crowell et al., 2009) and ingroup favoritism. The ingroup favoritism theory refers to humans' preference for the ingroup over the outgroup in terms of behavior, attitudes, preferences or perception (Turner et al., 1979). As the gender of the robot matches the ingroup of the participant, interaction with the social robot is likely to be evaluated more positively. This contrasts the other 50% of studies that found a positive effect for cross-matching robot gender and participant gender (Perugia & Lisy, 2022). Additionally, in some literature, there was an effect for male, but not for female participants (Crowell et al., 2009; Kuchenbrandt et al., 2014).

In general, literature suggests gendering social robots based on the gender associated with the stereotypical role they fulfill. Stereotypes are simplified representations of groups of people or things that are widely believed (Bryant et al., 2020). Social robots that matched their respective gender stereotypes were seen as more capable of performing behaviors of interest. Male robots were seen as more suitable for typically masculine tasks like guarding a house or repairing technical devices. Female robots were seen as more suitable for gender-stereotypically feminine tasks, such as household tasks and care (Eyssel & Hegel, 2012). Body shape can activate these stereotypes: people showed more affective trust towards a female robot torso than a male robot torso (Bernotat et al., 2017). Female virtual agents were perceived as more engaging, elicited more positive utility beliefs, and yielded higher math performances in an educational setting compared to a male virtual agent (Baylor, 2009; Plant et al., 2009). These results might be explained by the fact that, stereotypically, most schoolteachers are women. Powers et al. (2005) note that it may be beneficial to mismatch the gender to the stereotype on purpose, as humans adjust their speech to meet the perceived needs of others. For example, NurseBot (Carnegie Mellon University, United States) would stereotypically be gendered as female, because nursing is stereotypically gendered as a female profession. However, a male NurseBot may evoke more conversation, as patients might assume that a male NurseBot's knowledge is poorer than its female counterpart, which could help in acquiring more information from the patients.

It should be noted that using stereotypes in robot design and human-robot interaction must be done with caution to avoid the risk of reinforcing societal gender stereotypes. Weber (2005) suggested that robots are designed in the shape of women to make them appear as harmless and friendly, which would lead to more positive behaviour. However, Ghazali et al. (2018) argued that participants felt angrier and had more negative thoughts and beliefs towards a female robot advisor. According to Halberstam (1991), technology is gendered as female, as a female robot would be perceived as seductive and thus encourage engagement. UNESCO reported that current female digital assistants are too tolerant of mistreatment, are designed to be compliant and eager-to-please, are the representation of software errors, and are even mistaken for women in technology (West et al., 2019; Winkle et al., 2021). However, as Søraa (2017) points out, it is impossible to design genderless robots, as roboticists' understanding and ideas about the human form are inevitably influenced by a gendered perspective.

To answer the question in the title of this section: whether we should gender seemingly is context dependent. Matching genders might be beneficial, as the ingroup favoritism theory and social identity theory teach us. However, in practice cross-gendering might have a positive effect, as seen in Powers et al. (2005). At the same time, enforcing societal stereotypes on robots can be harmful. The contrasting views and mixed results in current literature are proof of the necessity of more research in this field. Therefore, in this study, the effect of gender on human-robot interaction is researched.

2.4. Human-robot interaction and persuasion

Persuasion is a fundamental aspect of social interaction. We can define persuasion as "any effort to modify an individual's evaluations of people, objects, or issues by the presentation of a message" (Petty & Cacioppo, 1986, p. 26). Persuasion in human-human interaction is widely researched, but whether the results of persuasion in HHI can be generalized to HRI is yet unknown. In this section, Aristotle's rhetorical triangle will be discussed, as his ideas formed the basis for persuasive strategies and the rhetorical triangle is relevant to the method of this research. Additionally, current results from persuasion in HRI are listed.

While often used synonymously, it is worth noting that persuasion and compliance are not the same thing. Compliance refers to yielding to a desire or demand from others, often someone in a position of power (Merriam-Webster, 2023). Persuasion can be used to achieve compliance. In this thesis, an attempt is made to persuade people of an opinion.

2.4.1. The rhetorical triangle

One of the first known attempts to define different persuasive strategies was done by Aristotle. He introduced his three modes of persuasion (Greek: *pisteis*) (Rapp, 2022): *ethos*, *logos* and *pathos*. These three modes of persuasion, also known as the rhetorical triangle, are still widely studied, and used today in various forms of communication including advertising (Vu, 2017), public speaking (Baccarani & Bonfanti, 2015; Haider, 2014) and writing (Isai et al., 2020). In this research, Aristotle's rhetorical triangle was used to distinguish between different strategies in persuasive speech.

Ethos (Greek for 'credibility') refers to the trustworthiness of the speaker or writer (Wu et al., 2022). It can be established through three aspects: practical intelligence, virtuous character, and good will. If a speaker or writer possesses all three of these aspects, there is no reason to doubt their credibility. However, ethos can be undermined if the speaker or writer lacks any of these three aspects. For example, if a speaker or writer has practical intelligence but lacks virtuous character or good will, the listener or reader may question their intentions. Similarly, if a speaker or writer has practical intelligence and a virtuous character but lacks good will, their advice may be seen as not being in the best interest of the audience. Lastly, if a speaker or writer lacks practical intelligence, virtuous character, and good will, their credibility and the value of their advice may questioned (Rapp, 2022).

Logos (Greek for 'word') refers to persuasion through the arguments of the speaker or writer (Wu et al., 2022). Arguments should be logically sound. There are two types of logical arguments. Inductive argumentation states that facts are determined by repeated observations. Deductive argumentation states that facts are obtained by deducing conclusions from premises. An example of a deductive argumentation scheme is a syllogism. Syllogisms often follow a structure with two premises and a conclusion. A famous example of a syllogism is:

"All men are mortal.
Socrates is a man.
Therefore, Socrates is mortal."

Aristotle notes that it does not matter if the syllogism is incorrect. The audience will take it to be the case that something has been proven, even if it is a fallacy. A competent rhetoric, however, is not fooled by incorrect arguments, as it should recognize fallacies and deception (Rapp, 2022).

Pathos (Greek for 'suffering' or 'experience') refers to the emotional impact that the speaker or writer has on its audience (Wu et al., 2022). The general idea is that the success of an argument depends on emotions, as we do not judge in the same way when we grieve, rejoice, or when we are happy. Emotion is defined as "all those feelings that so change men as to affect their judgements" (Aristotle in Roberts, 2004, p. 46). The speaker attempts to stir the hearers' emotions. However, pathos will only succeed if the emotion that the speaker wants to transfer to their audience is a strong emotion, such as anger, happiness or grief (Rapp, 2022).

2.4.2. Persuasive messages and current results from human-robot interaction

Persuasion in HHI and HRI is fundamentally different because social robots lack the capacity to be influenced or persuaded themselves. A robot should be considered social by the user, else, a robot will not be perceived as persuasive (Langedijk & Fischer, 2023). In their literature review of persuasive robots, Liu et al. (2022) distinguish between five factors that can influence the persuasiveness of the robot. The first factor is modality, where the presence of a social robot affected the persuasiveness of the robot positively. Interactivity is the second factor, where a robot was perceived as more persuasive when it was more interactive. Gestures negatively impacted persuasiveness. Third, results from tests of social character of a robot were less consistent. However, sociability and politeness affected the persuasiveness of a social robot. Higher sociability and providing negative feedback or touch led to higher levels of compliance. Fourth, literature showed that persuasive strategy yielded some effects, which suggests that there is potential to employ persuasive strategies from HHI in HRI. Lastly, context mattered: the more difficult a task was, the more persuasive the social robot, which was confirmed by Langedijk and Fischer (2023). Cultural background also impacted persuasiveness of the social robot. Fischer et al. (2019) showed that the speaking style of a robot affects persuasion. Steve Jobs' (considered a persuasive speaker) and Mark Zuckerberg's (considered a less persuasive speaker compared to Steve Jobs) speech characteristics were used. The results show that, even though the difference was not significant, prosodic characteristics of robot speech influenced the persuasiveness of the robot. Ruijten et al. (2015) found that social exclusion made people more susceptible to persuasion by an artificial agent. Additionally, men were less susceptible to the agent's feedback compared to women. This highlights that the effectiveness of persuasive technology can vary depending on the user's gender.

As briefly discussed in the introduction, there is no consensus in HRI on whether gender affects the persuasive abilities of a social robot. Additionally, studies vary in their methodologies and their definitions of persuasion. This makes it difficult to draw conclusions. In a literature review by Perugia & Lisy (2022), it was stated that persuasiveness had an effect in "two out of four studies" (p. 15). Even though literature on this specific topic is scarce, some studies that are related to robot gender and persuasion in HRI are reviewed.

Some literature found that robot gender had an effect on persuasion in HRI. Widely cited is a study done by Siegel et al. (2009). In their experiment, participants were given five dollars prior to the experiment as compensation for their participation. Then social robot Nexi (Massachusetts Institute of Technology, United States) asked them to donate money to the MIT Media Lab for research. The gender of robot Nexi was varied by changing the gender of the voice. A cross-gender preference was found: male participants preferred the female robot, female participants had little preference. This finding was based on the assumption that, when a robot was perceived as more trustworthy, credible, and engaging, the robot was more likely to change the subject's behavior, i.e., persuasion. Additionally, subjects donated more to the female robot than the male robot. Makenova et al. (2018) executed a similar experiment, using social robot NAO. Participants tended to donate more money to female robots compared to the male robot. Furthermore, foreigners donated more than locals, which is in line with the cultural background factor described by Liu et al. (2022).

Some studies found that robot gender had no effect on persuasion in HRI. In Ågren & Thunberg (2022), robot Furhat (Furhat Robotics, Sweden) gave a speech arguing that humans should not fear robots. The researchers measured persuasion using the rhetorical triangle. No significant difference in the robot's perceived persuasiveness was found based on gender. However, statistical evidence was found that male and female participants rated the robot's ethos differently. Female participants tended to rate robots more negatively compared to male participants. This is in line with Siegel et al. (2009). Thellman et al. (2018) varied in robot platform, as they used social robot NAO, but had a similar methodology. They reported no effect in perceived persuasiveness based on robot gender. Women rated the social robot as more persuasive than men, which contradicts Siegel et al. (2009), where men rated a female social robot as more persuasive than women. In Saunderson & Nejat (2022), two NAO robots were used. Participants had to guess how many jellybeans were in a jar. The robots provided suggestions using either pathos, logos, or no persuasive strategy. They found that the gender of the robot did not affect persuasive influence. However, for male participants, pathos was perceived as more persuasive than logos or the control condition.

To sum up, the results on the relationship between robot gender and a robots' persuasive abilities are inconsistent. Despite the lack of consensus, it is important to further explore this topic, as persuasion could play a role in HRI applications. This thesis aims to contribute to the existing body of literature by providing an answer to the question whether robot gender affects their persuasive capabilities.

2.5. Research question and hypotheses

As stated in the introduction, this thesis aims to find an answer to the following question: *How does the gender of a social robot's voice impact its ability to persuade humans?*

Based on the reviewed literature, we can draw some hypotheses. Based on Makenova et al.'s (2018) finding that participants donated more money to a female robot compared to a male robot, the first hypothesis is:

> **H1**: *Participants, regardless of their gender, are more persuaded by a female social robot compared to a male social robot.*

The second hypothesis of this thesis is formed based on Siegel et al.'s (2009) cross-gender preference:

> **H2**: *Male participants perceive a social robot as more persuasive compared to female participants.*

The third hypothesis is formed based on the ingroup favoritism theory proposed by Turner et al. (1979) and the social identity theory (Crowell et al., 2009):

> **H3**: *When the gender of a participant and the gender of a social robot match, participants perceive a social robot as more persuasive.*

Lastly, for the fourth hypothesis, a connection between papers is made. This is done as there is no known study that combines both persuasive strategy, participant gender and robot gender. In Saunderson & Nejat (2022) it is was found that emotion (i.e., pathos) was more effective for perceived persuasion than logic. Based on these results, and assuming H1 and H2, H4 is created:

> **H4:** *Male participants will perceive a female social robot that uses pathos as a persuasive strategy as most persuasive compared to persuasive strategies ethos and logos.*

Ch. 3 Method

To answer the research question of this thesis, three experiments were conducted: two pre-tests and one main experiment. After discussing the robot platform used in this research, the methods of the two pre-tests and their respective results are shown. The last section presents the method of the main experiment.

3.1. Robot Pepper

The social robot that is used in this thesis is robot Pepper. Robot Pepper is a semi-humanoid robot that has been designed by Softbank Robotics in 2014. Robot Pepper, in this study referred to as Pepper, is approximately 1.20 meters high. Peppers' design aims for a genderless look. However, user studies provide mixed results on Peppers' perceived gender. According to Seaborn and Frank's (2022) literature review, among the papers they analyzed, three papers perceived Pepper as female, while two papers identified Pepper as male and only one paper gendered Pepper as neutral. Additionally, only 12 papers, which represented 16% of the reviewed papers, discussed Peppers' gender. Participant attribution of Peppers' gender is also not widely described in literature; it was missing in 69% of the reviewed papers. In papers that did discuss gender attribution to Pepper by the participants, Pepper was labelled as male in 13 papers, as female in 15 papers and as neutral in 8 papers. Clearly, literature has not adopted the implied genderless look of Pepper. To eliminate gendering based on the appearance of the robot, the videos that were created for this experiment show Pepper from the waist up, as visible in Figure 1. The gender cue in this research is based on voice only. The tablet on Peppers' torso was not used in this research.

To program Pepper, a combination of software components was used: the Docker platform, the Social Interaction Cloud (SIC) framework, Google Dialogflow, and a custom Python SIC application. The SIC framework facilitated the connection to the robot Pepper through Docker. Docker is a tool that packages, provisions, and runs containers. Docker enabled access to applications that make use of Google Dialogflow. Google Dialogflow, Google's natural language understanding tool, was used to access the Google Cloud text-to-speech API on Pepper. The Python SIC application was used to generate the movements made by Pepper.

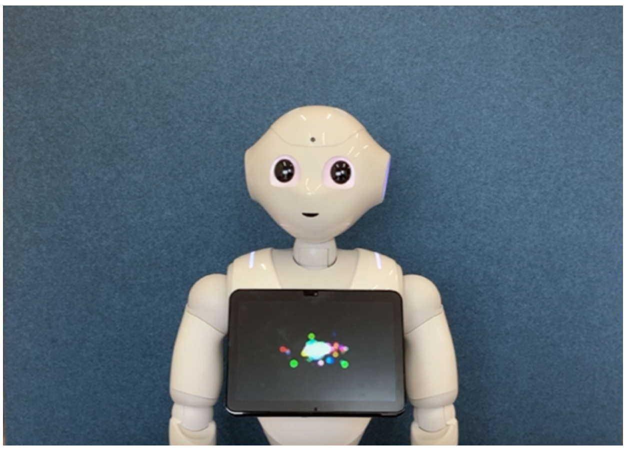

Figure 1. Screenshot of one of the videos that was created for this research. Pepper is shown from the waist up to eliminate gendering based on the appearance of the robot.

3.2. Pre-test 1

Goal of pre-test 1
Pre-test 1 was conducted to determine which voice Pepper would use in the main experiment. The goal of pre-test 1 was to ensure that the male robot voice was perceived as male, and that the female robot voice was perceived as female. In this way, pre-test 1 improved the accuracy of the main experiment.

Materials
The frequency of robot Peppers' voice was altered to create three male voices and three female voices. This was done using the website www.robotsindeklas.nl. The website was used to edit the voice pitch and speed of the voice. The voice pitch ranges from 0 to 200, where a lower number corresponds to a lower voice. Pitches higher than 100 were perceived as unpleasant by the researcher and were excluded from the pre-test. The six voices that were chosen thus ranged from 0 to 100 with increments of 20. Each voice spoke the same neutral sentence (adapted from Eyssel, Kuchenbrandt, Bobinger, et al. (2012)): *"*Good afternoon. According to my watch it is now a quarter past three. The train will leave in five minutes*."* Each audio recording took approximately 10 seconds.

Participants and procedure
A total of N = 20 participants rated robot voices; a within-subjects design was used. Participants were recruited using convenience sampling. A message with a link to the experiment was distributed via instant messaging service WhatsApp. Participants could complete the survey on any device.

Data was collected using the online survey platform Qualtrics. The survey took approximately five minutes to complete. Before starting the questionnaire, participants were asked to ensure that they could hear the audio files clearly, either by moving to a quiet room or by wearing headphones. Additionally, they were asked to evaluate the audio files independently. They could listen to the audio file multiple times if desired. Participants were informed that their participation was voluntary, that they could withdraw without providing a reason and that their anonymity was maintained throughout the whole experiment. Participants were not compensated for their participation.

No participants had to be excluded, as every participant provided their informed consent. All participants were aged 18 or older. Participants were asked whether they perceived the voice as more male, neutral, or more female on a 7-point scale. The order of the audio files was randomized. After answering the six questions in this survey, participants were thanked for their time and could end the survey.

3.3. Results pre-test 1

A total of 20 participants completed the survey. A Kruskal-Wallis H Test was conducted to examine whether the differences in gender perception were significant. Post hoc tests revealed a significant difference between the Voice 0 and Voice 100 (p < .001). Voice 80 and Voice 100 were perceived as female by all participants. However, none of the voices was perceived as male by all participants. Relatively, Voice 0 would be the best option: 12 out of

20 participants described the voice as male. However, eight participants perceived this voice as neutral or female. The ambiguity in the gender perception of the male voice led to the decision to not use the built-in voice of robot Pepper.

As the voice of the robot is one of the manipulated factors in the main experiment, it is essential that the voice was perceived correctly according to its gender. For this research, the Google Cloud text-to-speech API was used. Using audio files instead of the robots' voice is common in HRI. For example, Pourfannan et al. (2022) used Notevibes AI voice generator, Kuchenbrandt et al. (2014) employed Acapela Mobility 7.0, and Chita-Tegmark et al. (2020) used the Mac OS text-to-speech voices Alex and Samantha. The female voice en-US-Neural2-F and male voice en-US-Neural2-J were selected in this study. Both voices were perceived as having a clear and neutral tone by the researcher and are proven to be correctly recognized as either male or female. Additionally, the voices were identical in intonation, accentuation, and rhythm. The Google Cloud text-to-speech API was employed to generate text to synthetic voice audio files. These audio files were edited to synchronize with the videos.

3.4. Pre-test 2

Goal of pre-test 2
Three types of persuasive strategies are used in this research, namely ethos, pathos, and logos (for definitions, see section 2.4.1). To ensure that the speech used in the main experiment reflects the values of ethos, logos, and pathos sufficiently, the speeches were evaluated beforehand.

Materials
The Persuasive Discourse Inventory (PDI) (Feltham, 1994) is used for pre-test 2, similar as in Thellman et al. (2018) and Ågren & Thunberg (2022). The PDI scale comprises 17 questions distributed over ethos, pathos, and logos.

Three speeches were created, in which robot Pepper tries to convince the reader or listener that there is no reason to fear robots, similar to Thellman et al. (2018) and Ågren & Thunberg (2022). All three speeches had a baseline in their textual flow. However, the arguments used to convince the participant differentiated according to their respective persuasive strategy, i.e., ethos, logos, or pathos. The arguments used in the speeches were created based on the scales that are used in the PDI. This worked in the following way. The first scenario aimed to reflect values of ethos. In the PDI scale, ethos is described as "*believable*", "*credible*", "*trustworthy*", "*reliable*", and "*dependable*". The second scenario aimed to reflect values of logos. In the PDI scale, logos is described as "*rational*", "*informative*", "*deals with facts*", "*knowledgeable*", and "*logical*". The third scenario aimed to reflect the values of pathos. The PDI scale describes pathos as "*stirring*", "*moving*", "*exciting*", "*stimulating*", "*reaches out to me*", "*affects my feelings*", and "*touches me emotionally*". The ethos scenario had 304 words, the logos scenario had 301 words, and the pathos scenario had 305 words. The pre-tested speeches are included in Appendix A.

Participants and procedure
A between-subjects design was used. Participants could complete the survey on any device. A total of 50 participants (M_{age}= 27.73, SD = 8.67; 25 women, 25 men) were recruited using the platform Prolific. Participants were financially compensated for completing the survey.

The survey was created using the online survey platform Qualtrics. The survey took approximately eight minutes to complete. The survey was shared with another researcher; thus the survey was divided in two parts. The results from the other researcher are not considered in this thesis. Each participant filled out both parts.

Before starting the questionnaire, participants were informed that their participation was voluntary, that they could withdraw without providing a reason and that anonymity was maintained throughout the whole experiment. Participants were recruited using the platform Prolific. Participants were financially compensated for their participation. No participants had to be excluded, as every participant provided their informed consent. All participants were aged 18 or older.

Participants were instructed that they would read one written scenario in which a robot gives a speech and that they would be asked a total of 17 questions about their perception of the robot after reading the scenario. The scenario remained visible when filling in the questions. Participants read one written scenario: either the ethos, pathos, or the logos scenario. Each question followed the same format: participants had to tick the box that they felt best described the scenario they read on a semantic 7-point scale. The order of the questions was randomized. After answering all questions in this survey, participants were thanked for their time and could click the button to be redirected to Prolific which would register their submission and end the survey.

3.5. Results pre-test 2

A total of 50 participants completed the PDI (ethos: N = 16, logos: N = 17, pathos: N = 17). Three one-way ANOVAs were performed to compare whether the results of scenario 1, scenario 2 and scenario 3 reflected their intended persuasive strategy. Note that that the assumption of normality for a one-way ANOVA was violated. Nevertheless, statistical test ANOVA was used: the ideal PDI will yield skewed results, as it is a semantic scale as opposed to a Likert scale.

Scenario 1: ethos. The first scenario aimed to reflect the values of ethos. The one-way ANOVA for scenario 1 revealed that there was a statistically significant difference in the mean scores for ethos, pathos and logos, where $F(2, 45) = 14.45$, $p < .001$. Post-hoc comparisons using the Tukey HSD Test indicated that the mean score for ethos (M = 24, SD = 4.95) was significantly higher than the pathos condition (M = 16.87, SD = 5.20). The logos condition (M = 26.43, SD = 5.51) did not significantly differ from the ethos condition. A comparison between logos and pathos was not necessary, as it would not impact the results, as this was not the purpose of the scenario.

As the ethos condition did not differ significantly from the logos condition, scenario 1 was edited to ensure that it did not reflect logos instead of ethos.

Scenario 2: logos. The second scenario aimed to reflect the values of logos. The one-way ANOVA for scenario 2 revealed that there was a statistically significant difference in the mean scores for ethos, pathos and logos, where $F(2, 48) = 17.42$, $p < .001$. Post-hoc comparisons using the Tukey HSD Test indicated that the mean score for logos ($M = 28.41$, $SD = 3.96$) was significantly higher than the ethos condition ($M = 23.71$, $SD = 5.53$) and the pathos condition ($M = 18.91$, $SD = 4.46$). A comparison between ethos and pathos was not necessary, as it would not impact the results, as this was not the purpose of the scenario.

Given these results, we concluded that Scenario 2 correctly reflected the values of logos.

Scenario 3: pathos. The third scenario aimed to reflect the values of pathos. The one-way ANOVA for scenario 3 revealed that there was a statistically significant difference in the mean scores for ethos, pathos and logos, where $F(2, 48) = 6.79$, $p = .003$. Post-hoc comparisons using the Tukey HSD Test indicated that the mean score for pathos ($M = 16.72$, $SD = 6.38$) was significantly lower than the logos condition ($M = 24.76$, $SD = 5.90$). The ethos condition ($M = 19.24$, $SD = 7.17$) did not differ significantly from the pathos condition. A comparison between ethos and logos was not necessary, as it would not impact the results, as this was not the purpose of the scenario.

Based on these results, scenario 3 was edited as a whole: the mean of pathos is lower than the mean of logos, indicating that scenario 3 reflected the values of logos more than those of pathos, while the intent of scenario 3 was to reflect pathos. Additionally, ethos did not differ significantly from pathos.

The edited speeches that were used in the main experiment are included in Appendix B.

3.6. Main experiment

The main experiment was designed to answer the research question of this thesis. An online, video-based experiment was designed, in which two measures were varied: the gender of the robot, as well as persuasive strategy used in the speech. The Ethics and Privacy Quick Scan of the Utrecht University Research Institute of Information and Computing Sciences classified this research as low risk with no fuller ethics review or privacy assessment required.

Design

The main experiment had a 3 (persuasive message: ethos versus pathos versus logos) x 2 (gender participant: male versus female) x 2 (gender robot: male voice versus female voice) between-subjects design. A power analysis was performed to calculate the sample size *a priori*, using the tool G*Power, with statistical test ANOVA (fixed effects, special, main effects, and interaction). The sample size for a design with 12 groups needs to consist of at least 206 participants to be able to detect a medium effect size ($f = .25$) with a standard power ($1-\beta = .9$), 2 degrees of freedom[2], and standard error probability ($\alpha = .05$). The

[2] Degrees of freedom was calculated using the following formula: $df = (n_1 - 1) \times (n_2 - 1) \times (n_3 - 1)$, as this research examines a possible interaction effect between 3 (persuasive message: ethos versus pathos versus logos) x 2 (gender participant: male versus female) x 2 (gender robot: male voice versus female voice). The calculation for the degrees of freedom is thus as follows: $df = (3 - 1) \times (2 - 1) \times (2 - 1) = 2$.

values used here are similar to Saunderson & Nejat (2022), except for a higher standard power.

Participants

A total of N = 250 participants were recruited using the platform Prolific. Participants were financially compensated for completing the survey. Ten participants were excluded as they did not fill out the entire questionnaire, resulting in N = 239. The gender distribution was 117 women, 115 men, four non-binary participants, two preferred not to say and one preferred to self-describe. Unfortunately, the group of non-binary participants was not large enough to hold sufficient statistical power. Any participant that did not identify as male or female, or preferred not to disclose their gender was excluded from analysis. Although we acknowledge that this decision reflects a simplified understanding of gender, the group was not large enough to hold sufficient statistical power. One participant failed the attention check and was thus excluded. This reduced the total amount of participants to N = 231 participants included for data analysis, with 116 women and 115 men. This means that the required sample size of N = 206 was met. Participants were between 18 and 66 years old (M_{age}= 29.01, SD = 10.01).

Two questions were added in the questionnaire to assess the level of knowledge participants had on robots. 30 participants (12.9%) reported to be not at all knowledgeable about robots, 90 participants were somewhat knowledgeable (39.2%), 82 participants were moderately knowledgeable (35.3%). 11.6% and .9% reported to be respectively very knowledgeable and extremely knowledgeable. As to how often the participants had encountered robots in the last year, 23.7% did not, 42.2% rarely, 27.2% sometimes, 6.5% often, and .4% encountered robots all the time.

Procedure

The survey was created using the online survey platform Qualtrics. Completing the survey took approximately 11 minutes (M = 640.54 seconds, SD = 273.22).

Participants were first asked to read the informed consent. The informed consent stated that participation is voluntary, and the research was completely anonymous. All participants were aged 18 years or older. The researchers' email address and the academic integrity counsellor's email address and telephone number were added in the informed consent, in case of questions about rights as a participant. Participants could complete the survey on any device but were informed that this survey was optimized for display on a (desktop) computer. Participants could agree or not agree, in case of the latter participants were immediately directed to the end of the experiment.

The survey contained three questionnaires. The first questionnaire that was administered was the GAToRS questionnaire. The order of the questions was randomized. Then, an introductory video of social robot Pepper was presented. The video was uploaded to YouTube and had been embedded in the Qualtrics survey. This video was added to address the novelty effect. The novelty effect is described as a positive effect related to the newness of a situation rather than the situation itself, that wears off over time (Elston, 2021). With the introductory video, participants could get used to the kind of movements that robot

Pepper would make (e.g., waving). The video contained no sound and had a duration of 43 seconds.

Next, participants were directed to one of six videos. The videos had been uploaded to YouTube and had been embedded in the Qualtrics survey. The video began with a disclaimer that this video would contain sound, and to ensure that the participant could hear the video properly. Each participant watched one video, encapsulating the between-subjects design. Two aspects of the videos were varied: the gender of the robot voice (either male or female), and the persuasive message (using either ethos, logos, or pathos as a persuasive strategy). Robot Pepper had either a male or female voice (as validated in pre-test 1) and gave one of the three speeches that was validated in pre-test 2. Each video had an equal number of words (N = 322); however, some words took longer to pronounce. As a result, videos varied in length, ranging from 139 seconds to 149 seconds. Each video ended with a message that they could return to the questionnaire. The videos contained subtitles to improve the accessibility of the survey.

One multiple choice question was added after participants watched the video as an attention check. Then, participants filled out the second questionnaire: the RoSAS questionnaire (Carpinella et al., 2017). The third questionnaire that was administered was a questionnaire on the perceived persuasiveness of the message (Mullennix et al., 2003). The orders of the questions of both questionnaires were randomized.

Then, a randomized version of the GAToRS questionnaire was asked again. This was done in order to compare the opinions of the participants before exposure to the video and after exposure to the video. A difference in the scores on the first and second survey would indicate a change in opinion and would thus suggest persuasive influence of the robot. Two demographic questions were included. After completing all questionnaires, participants were shown a debriefing that contained the goals and expectations of this research. Participants were then thanked for their participation and redirected to the Prolific platform.

In total, the questionnaire consisted of 55 questions. The questionnaire flow is shown in Figure 2.

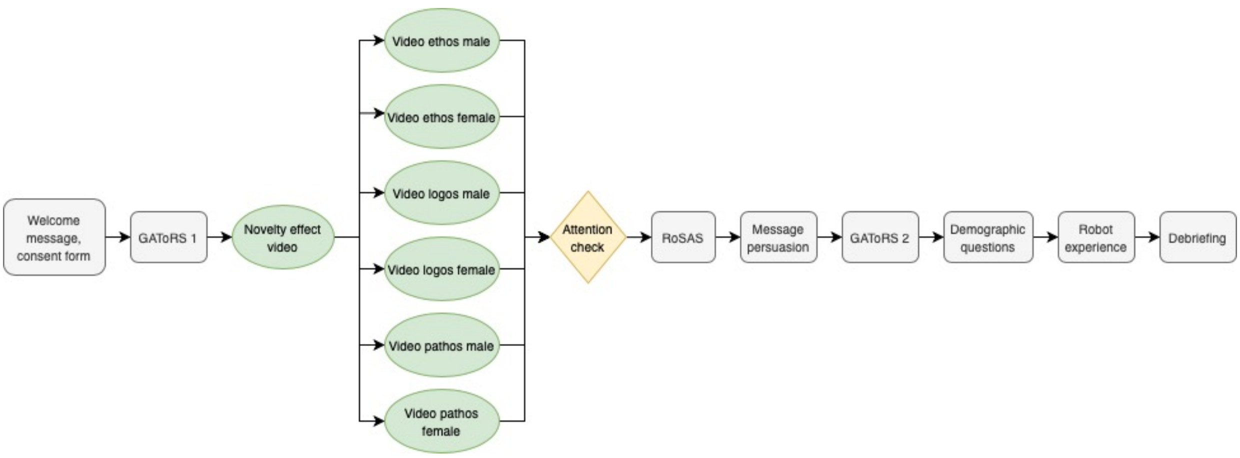

Figure 2. A visual representation of the order of the surveys in the questionnaire. Three questionnaires were used: GAToRS, RoSAS, and a questionnaire on message persuasion.

Instruments

***GAToRS questionnaire*.** The GAToRS questionnaire is the first questionnaire that was presented to the participants. The GAToRS questionnaire is relatively new, and was designed to be an instrument to measure general multidimensional attitudes towards robots (Koverola et al., 2022). In this research, two of the four subscales of the GAToRS questionnaire were used, namely the societal level positive (S+) and societal level negative (S-), as this research focuses on general attitudes towards robots, not personal opinions. This resulted in a total of ten questions (five questions for S+; five questions for S-). Participants were asked to answer the questions on a 7-point scale (1 = "*strongly disagree*"; 7 = "*strongly agree*"), keeping their general opinion on robots in mind.

***RoSAS questionnaire*.** The RoSAS questionnaire (Carpinella et al., 2017) is a well-known questionnaire in HRI that is used to measure people's judgments of the social attributes of robots. The RoSAS questionnaire is divided in three dimensions: competence, warmth, and discomfort as shown in Figure 3. Cronbach's alpha for each dimension was α = .838, α = .838 and α = .828 respectively. There are six questions per dimension, resulting in a total of eighteen questions. Participants were asked to fill out their views on a 7-point scale regarding the robot in the video, ranging from "*not at all*" to "*very much so*".

Competence	Warmth	Discomfort
Reliable	Organic	Awkward
Competent	Sociable	Scary
Knowledgeable	Emotional	Strange
Interactive	Compassionate	Awful
Responsive	Happy	Dangerous
Capable	Feeling	Aggressive

Figure 3. The three dimensions of the RoSAS questionnaire (Carpinella et al., 2017).

***Perceived persuasiveness of the message*.** The third questionnaire that was used contains two sections of a questionnaire that measures the perceived persuasiveness of the message (Mullennix et al., 2003). This questionnaire is used to test whether the perceived persuasiveness of the message has an impact on a possible change in opinion. The two subscales that were used were six items on the perceptions of the message ("*stimulating-boring*"; "*vague-specific*"; "*unsupported-supported*"; "*complex-simple*"; "*convincing-unconvincing*"; "*uninteresting-interesting*") and six items on the perceptions of the effectiveness of the argument ("*bad–good*"; "*foolish–wise*"; "*negative–positive*"; "*beneficial–harmful*"; "*effective–ineffective*"; "*convincing–unconvincing*") on a 7-point scale. Cronbach's alpha for each dimension was α =.730 and α = .928 respectively.

***Attention check*.** One question was added as an attention check after watching the second video. This was done to ensure that the participants had watched the video and had listened to the content. Data entered by participants who failed this attention check were perceived as invalid for this research. The following question had to be answered: "*What was the subject the robot was talking about?*". Possible answers were: "*People should not be afraid of robots*."; "*Robots are good teachers*."; "*The robot talks about geography*".

***Demographics*.** In addition to age, the gender of the participant was a relevant factor. To ensure an inclusive approach for all, this question was designed with Spiel et al.'s (2019) guidelines in mind. Five options were given: "*woman*"; "*man*"; "*non-binary*"; "*prefer* *to not disclose*"; and "*prefer to self-describe*". The last option contained a blank text box. Lastly, two additional questions were included to assess the participants' level of knowledge regarding robots. The first question requested an indication of how knowledgeable participants were about the robot, ranging from "*not at all knowledgeable*" to "*extremely knowledgeable*", on a 5-point scale. The second question asked how often participants had encountered robots in the last year (on a scale of "*never*" to "*all the time*" on a 5-point scale).

Ch. 4 Results

This results-section has been split up in multiple sections. First, the results of the attention check are discussed, as well as the setup of the data analysis. Then, for each hypothesis, the expected interaction effects are investigated. Lastly, a short analysis of the covariates is presented.

Attention check
As described in section 3.6, this experiment contained a question to ensure that participants had viewed the video (attention check). One participant was excluded from the dataset as they answered the attention check question incorrectly. This reduced the number of participants to $N = 231$, with 116 female participants and 115 male participants. Participants generally performed well on the attention check question, as 99.6% of the participants answered correctly.

Setup of data analysis
Results were analyzed using IBM SPSS Statistics 28. As the GAToRS questionnaire is separated in S+ and S-, they have been evaluated separately. In other words, every hypothesis that uses the GAToRS questionnaire was evaluated twice: once for S+ and once for S-. This research aimed to examine whether a change in the participants' opinion has occurred due to the manipulation. This was done by comparing the results of the first and second time the participants completed the questionnaire. Two variables were created to be able to analyze the results:

- score_positive, which was calculated by subtracting the answers that were given in the second S+ GAToRS questionnaire by the first S+ GAToRS questionnaire. The score ranged between -7 and 7. Cronbach's alpha for the first and second S+ GAToRS questionnaire was $\alpha = .738$ and $\alpha = .835$ respectively.
- score_negative, which was calculated by subtracting the answers that were given in the second S- GAToRS questionnaire by the first S- GAToRS questionnaire. The score ranged between -7 and 7. Cronbach's alpha for the first and second S- GAToRS questionnaire was $\alpha = .751$ and $\alpha = .781$ respectively.

The RoSAS questionnaire and message persuasion questionnaire are used as covariates. This is done to account for the influence of some confounding variables. The RoSAS questionnaire is distributed over three dimensions: competence, warmth, and discomfort. Each dimension was added as a separate covariate; whether a person perceived a robot as competent, warm, or whether the robot made the person feel uncomfortable could have impacted perceptions of the robot. The message persuasion questionnaire that was used also consists of two subscales: one on perceptions of the message and one on the perceptions of the effectiveness of the argument. Each subscale was added as a separate covariate; whether the participant perceived the message as e.g., vague, or too simple, could have impacted opinions. It should be noted that other variables could also have affected the results. However, it is impossible to account for all external variables.

Before testing the hypotheses, the assumptions for two two-way ANCOVA between groups were tested. The Shapiro Wilk test of Normality was used to assess the assumption of normality (score_positive: $p < .001$, score_negative: $p < .001$). Bar graphs showing the normal distribution curve of score_positive and score_negative are included in Appendix C. Levene's test of Equality of Error Variances was used to test the assumption of homogeneity of variance. For score_negative, this assumption was met (p = .788). However, for score_positive, the assumption was violated (p = .002). According to Field (2016), this does not pose a significant challenge. All covariates were linearly related to the dependent variable. The normality of the residuals was evaluated, and this assumption was met. Lastly, the homogeneity of regression slopes was tested. For score_positive, a significant value was found for covariate RoSAS discomfort and the gender of the robot (p = .018). For this reason, the RoSAS discomfort scale was dropped as a covariate for the ANCOVAs in all hypotheses for score_positive.

An overview of the results for the main effects and all interactions for score_positive and score_negative is included in Appendix D.

H1: *Participants, regardless of their gender, are more persuaded by a female social robot compared to a male social robot.*

To test hypothesis 1, the main effect between the gender of the robot and perceived persuasiveness was reviewed. Two one-way repeated measures ANCOVAs were conducted, where the gender of the robot was the independent variable and the test results to the GAToRS questionnaire were the dependent variables. For score_positive, values are corrected for covariates RoSAS warmth, RoSAS competence, persuasiveness of the message and persuasiveness of the argument. For score_negative, values are corrected for covariates RoSAS warmth, RoSAS competence, RoSAS discomfort, persuasiveness of the message, and persuasiveness of the argument.

A significant main effect between the gender of the robot and perceived persuasiveness was found for score_positive, with $F(1, 215) = 4.505$, $p = .035$. For score_negative, no significant main effect between the gender of the robot and perceived persuasiveness was found, $F(2, 214) = .022$, $p = .978$. Post-hoc comparisons for score_positive using the Bonferroni correction indicated that the mean score on GAToRS societal positive for participants that saw a female robot ($M = .279$, $SD = .055$) were significantly higher than for participants who saw a male robot ($M = .113$, $SD = .056$). These results indicate that participants who saw a female robot were more affected in their positive attitude towards robots than participants who saw a male robot in the manipulation. The female robot ($M = -.328$, $SD = .632$) was not perceived as more persuasive compared to the male robot ($M = -.325$, $SD = .759$) for score_negative.

When looking at the actual scores on the GAToRS questionnaire of the pre- and post-test in Figure 4, we see a relevant, but as discussed, not significant tendency where the score for S+ increases faster in the condition with the female robot than the male robot. For S- this effect is less visible. Participants were positively affected in their opinions on robots for both score_positive and score_negative: for score_positive, the mean score increased, and for score_negative the mean score decreased. We will thus partially accept H1.

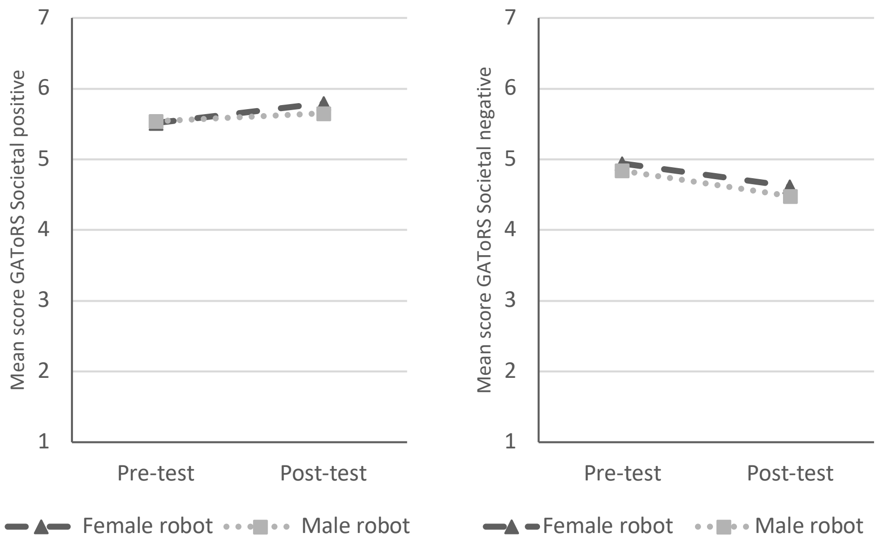

Figure 4. The mean results of the pre-test GAToRS scores compared to the post-test GAToRS scores for both societal level positive (left) and societal level negative (right), for participants who saw either a male or a female robot, on a scale of 1 to 7. A significant main effect was found for societal level positive (p = .022). For S+, the means are adjusted for covariates warmth, competence, perceived persuasiveness of the message and persuasiveness of the argument. For S-, the means are adjusted for covariates warmth, competence, discomfort, perceived persuasiveness of the message and persuasiveness of the argument.

H2: *Male participants perceive a social robot as more persuasive compared to female participants.*

The main effect between the gender of the participant and perceived persuasiveness was reviewed for hypothesis 2. Two one-way repeated measures ANCOVAs were conducted, where the gender of the participant was the independent variable and the test results to the GAToRS questionnaire were the dependent variables. For score_positive, values are corrected for covariates RoSAS warmth, RoSAS competence, persuasiveness of the message and persuasiveness of the argument. For score_negative, values are corrected for covariates RoSAS warmth, RoSAS competence, RoSAS discomfort, persuasiveness of the message and persuasiveness of the argument.

No significant main effect between the gender of the participant and perceived persuasiveness was found for score_positive, with $F(2, 215) = 1.464$, $p = .228$. A significant main effect between the gender of the participant and perceived persuasiveness was found for score_negative, with $F(2, 214) = 8.237$, $p = .005$. Post-hoc comparisons using the Bonferroni correction indicated that the mean score for female participants (M = -.462, SD = .060) was significantly higher than for male participants (M = -.217, SD = .060). This tendency is visible in Figure 5. These results suggest that female participants were more affected by the manipulation on score_negative than male participants. In other words, female participants had a less negative attitude towards robots after the manipulation compared to male participants, regardless of persuasive message or robot gender. For score_positive, male participants (M = .141, SD = .488) did not perceive the social robot as more persuasive compared to female participants (M = .252, SD = .708). Both male and female participants responded more negatively on the second GAToRS questionnaire. This implies that participants disagreed more with the negative statements in the second GAToRS societal negative, thus suggesting that participants were somewhat persuaded by the social robot.

For score_negative, the mean score in the second GAToRS questionnaire is *M* = 4.565 (*SD* = 1.199) for all participants. As this value is over 4 on a 7-point scale, it is assumed that participants generally agreed with the negative statements on robots. The manipulation was thus not successful in bringing the mean score down to a value lower than 4, which would have indicated that participants would have become overall not scared of robots.

We cannot accept H2, as H2 should have yielded significant results for both score_positive and score_negative. Additionally, female participants scored significantly lower on the second GAToRS questionnaire compared to male participants, which contradicts the hypothesis. Thus, we reject H2.

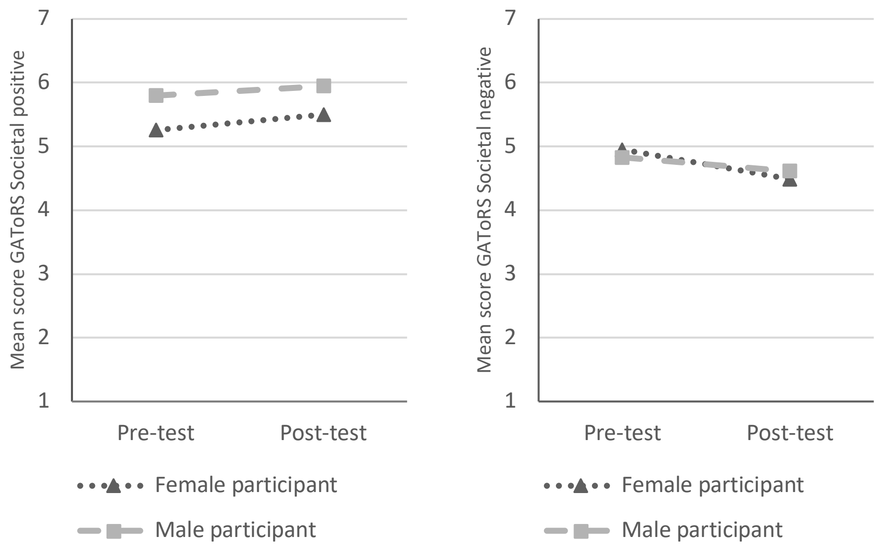

Figure 5. The mean results on the pre-test GAToRS score compared to the post-test GAToRS score for societal level positive (left) and societal level negative (right) for male versus female participants, on a scale of 1 to 7. A significant main effect was found for societal level negative (*p* = .005). For S+, means are adjusted for covariates warmth, competence, perceived persuasiveness of the message and persuasiveness of the argument. For S-, means are adjusted for covariates warmth, competence, discomfort, perceived persuasiveness of the message and persuasiveness of the argument.

H3: *When the gender of a participant and the gender of a social robot match, participants perceive a social robot as more persuasive.*

To test hypothesis 3, the interaction effect between the gender of the robot, the gender of the participant and perceived persuasiveness was reviewed. A gender-match is described in this research as a female participant who saw a female robot, and a male participant who saw a male robot (*N* = 127). 104 participants saw a robot with a different gender than themselves. Two two-way repeated measures ANCOVAs were conducted, where the gender of the robot and the gender of the participant were independent variables and the test results to the GAToRS questionnaire were the dependent variables. For score_positive, values are corrected for covariates RoSAS warmth, RoSAS competence, persuasiveness of the message and persuasiveness of the argument. For score_negative, values are corrected for covariates RoSAS warmth, RoSAS competence, RoSAS discomfort, persuasiveness of the message and persuasiveness of the argument.

No significant interaction effect between a gender-match of the participant and the robot and perceived persuasiveness was found, with score_positive: $F(1, 215) = .038$, $p = .846$, and score_negative: $F(1, 214) = 1.968$, $p = .162$. Participants who saw a robot with the same gender as them (female participant-female robot: $M = .325$, $SD = .530$, male participant-male robot: $M = .083$, $SD = .513$) did not perceive the robot as more persuasive compared to participant who saw a robot with a different gender (female participant-male robot: $M = .162$, $SD = .877$, male participant-female robot: $M = .212$, $SD = .452$) for score_positive. For score_negative, participants who saw a robot with the same gender as them (female participant-female robot: $M = -.388$, $SD = .626$, male participant-male robot: $M = -.181$, $SD = .623$) also did not perceive the robot as more persuasive compared to participant who saw a robot with a different gender (female participant-male robot: $M = -.500$, $SD = .872$, male participant-female robot: $M = -.254$, $SD = .638$). A gender-match between the participant and the social robot thus had no significant effect on the perceived persuasiveness of the robot. Therefore, hypothesis 3 is rejected.

H4: *Male participants will perceive a female social robot that uses pathos as a persuasive strategy as most persuasive compared to persuasive strategies ethos and logos.*

For hypothesis 4, the interaction effect between the gender of the participant, the gender of the robot, and perceived persuasiveness was reviewed. Two three-way repeated measures ANCOVAs were conducted, where the gender of the robot is an independent variable, the gender of the participant is the second independent variable, and the persuasive message is the third independent variable. The dependent variable is the result on the GAToRS questionnaire, i.e., either score_positive or score_negative. For score_positive, values are corrected for covariates RoSAS warmth, RoSAS competence, persuasiveness of the message and persuasiveness of the argument. For score_negative, values are corrected for covariates RoSAS warmth, RoSAS competence, RoSAS discomfort, persuasiveness of the message and persuasiveness of the argument.

No significant interaction was found between the gender of the participant, the gender of the robot, and perceived persuasiveness, with score_positive: $F(2, 215) = .120$, $p = .887$, score_negative: $F(2, 214) = 1.624$, $p = .199$. For score_positive, male participants that saw a female social robot that uses pathos as a persuasive strategy ($M = .227$, $SD = .518$) did not perceive the robot as more persuasive compared to persuasive strategies ethos ($M = .211$, $SD = .397$), logos ($M = .200$, $SD = .475$), and participants who saw a male robot (ethos: $M = .000$, $SD = .624$; logos: $M = .152$, $SD = .510$; pathos: $M = .080$, $SD = .440$). The means and standard deviations are displayed in Table 1. For score_negative, male participants who saw a female social robot that used pathos as a persuasive strategy ($M = -.400$, $SD = .809$) did not perceive the robot as more persuasive compared to persuasive strategies ethos ($M = -.126$, $SD = .612$), logos ($M = -.240$, $SD = .394$), and participants who saw a male robot (ethos: $M = -.329$, $SD = .565$; logos: $M = -.324$, $SD = .592$; pathos: $M = .040$, $SD = .643$). The means and standard deviations are displayed in Table 2. This means that this research found no evidence that male participants perceive a female social robot that uses pathos as a persuasive strategy as most persuasive compared to persuasive strategies ethos and logos. Thus, hypothesis 4 is rejected.

Certain observations can be derived from Table 1. Higher means were found on the GAToRS societal positive questionnaire for female participants than male participants, except for the male robot – pathos condition. This suggests a (non-significant) tendency where women are more persuaded by the manipulation. Additionally, for the condition male participant – male robot – ethos, a mean of *M* = .000 was found, which indicates that participants did not change their opinion after the manipulation. Table 2 also reveals some interesting findings. Pathos seems most effective as a persuasive strategy to change negative opinions to more positive opinions. However, this observation does not hold for male participants that saw a male robot: here we see the only positive mean for score_negative (*M* = .040). This suggests a (non-signifianct) tendency in which male participants that saw a male robot in the condition pathos had a more negative opinion after watching the video. A last observation from Table 1 and Table 2 is that persuasive strategy ethos yielded higher results for female participants compared to male participants. However, none of these observations were significant, as the ANCOVA for both score_positive and score_negative was not significant. These observations only reflect inclinations and should therefore be approached with caution.

Table 1

Means and standard deviations for score_positive for 3 (persuasive message: ethos versus pathos versus logos) x 2 (gender participant: male versus female) x 2 (gender robot: male voice versus female voice).

		Female Participant		Male Participant	
		M	*SD*	*M*	*SD*
Female robot					
	Ethos	.368	.500	.211	.397
	Logos	.256	.512	.200	.475
	Pathos	.370	.592	.227	.518
Male robot					
	Ethos	.286	.763	.000	.625
	Logos	.157	.741	.152	.510
	Pathos	.012	1.110	.080	.440

Note. The values are adjusted for covariates RoSAS warmth, RoSAS competence, persuasiveness of the message and persuasiveness of the argument. N = 231.

Table 2

Means and standard deviations for score_negative for 3 (persuasive message: ethos versus pathos versus logos) x 2 (gender participant: male versus female) x 2 (gender robot: male voice versus female voice).

		Female Participant		Male Participant	
		M	*SD*	*M*	*SD*
Female robot					
	Ethos	-.347	.696	-.126	.612
	Logos	-.384	.568	-.400	.809
	Pathos	-.430	.656	-.240	.394
Male robot					
	Ethos	-.384	.568	-.329	.565
	Logos	-.500	.917	-.324	.592
	Pathos	-.612	.757	.040	.643

Note. The values are adjusted for covariates RoSAS warmth, RoSAS competence, RoSAS discomfort, persuasiveness of the message and persuasiveness of the argument. N = 231.

Analysis of covariates

The RoSAS questionnaire is used to measure people's judgements of the social attributes of robots, separated in competence, warmth, and discomfort. This questionnaire is not used as a means to answer the hypotheses in this research, as this questionnaire was presented only once, which makes it impossible to measure a difference in opinion before and after manipulation. However, some interesting observations can be made. The mean scores on each of the subscales for RoSAS are shown in Figure 6. The first observation is the gender-match effect on RoSAS competence. Female participants rated the female robot as more competent (M = 5.037, SD = .785) compared to the male robot (M = 4.708, SD = 1.080). Male participants rated the male robot as more competent (M = 4.905, SD = .931) compared to the female robot (M = 4.747, SD = 1.011). A gender-match between the robot and the participant could thus potentially have a positive effect on ratings for competence. The second observation is on RoSAS warmth. Male participants gave higher ratings to warmth for both male robots (M = 3.643, SD = 1.023) and female social robots (M = 3.664, SD = 1.094) compared to female participants (female robot: M = 3.407, SD = 1.094; male robot: M = 3.301, SD = 1.238). Lastly, both the female and male robot score low on discomfort (female robot: M = 2.211, SD = .917; male robot: M = 2.309, SD = 1.117). This is a desirable outcome: participants did not feel too uncomfortable watching the video.

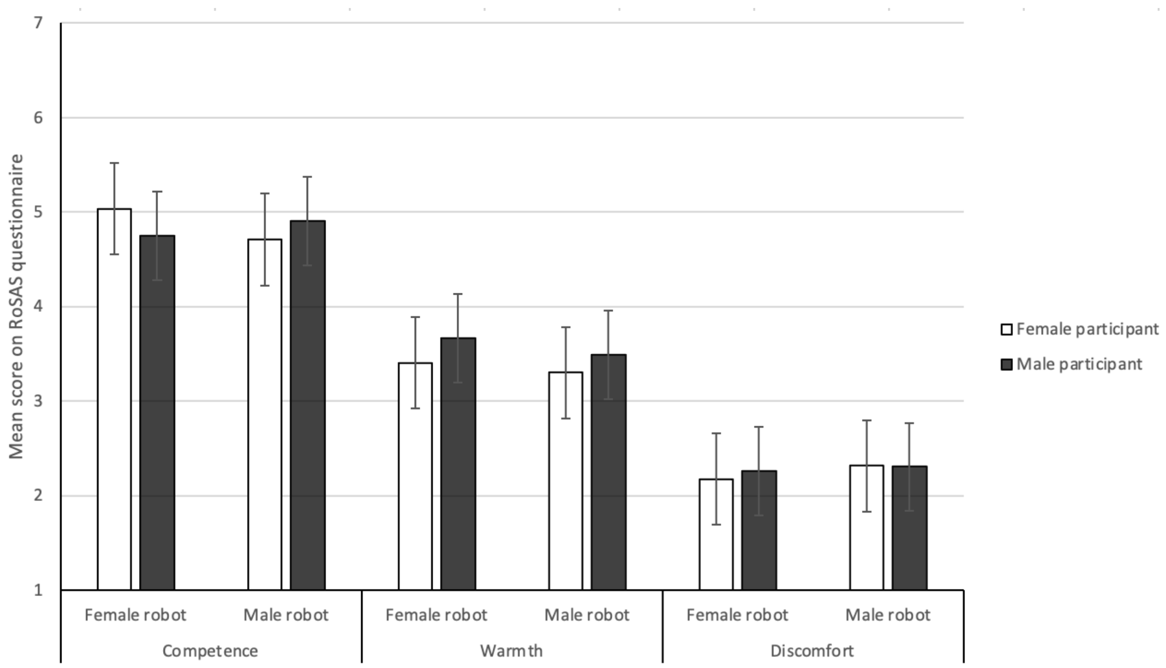

Figure 6. The mean scores of the RoSAS questionnaire, separated on scales competence, warmth, and discomfort and separated on the gender of the robot. The error bars represent standard errors.

Secondly, two scales from Mullenix et al.'s (2003) questionnaire were used to test perceived persuasiveness. This questionnaire asked participants to evaluate the message conveyed by the robot in the video. The mean scores for all scales are over 4 on a scale of 1 to 7. We can thus assume that participants generally thought that the speech contained the correct elements to be perceived as persuasive. One observation can be made from Figure 7. For the scale effectiveness of the argument, we see that female participants gave the female robot on average higher ratings (M = 5.323, SD = 1.221) compared to the male robot (M = 5.276, SD = 1.351). Male participants gave the male robot higher ratings (M = 5.471, SD = 1.300) compared to the female robot (M = 5.244, SD = 1.285). We find a similar tendency for perceptions on the message. Female participants gave the female robot higher ratings (M = 5.193, SD = .773) compared to the male robot (M = 5.016, SD = 1.041), male participants gave the male robot higher ratings (M = 5.071, SD = 1.066) compared to the female robot (M = 4.965, SD = .902). This finding is peculiar, as both the male and female robot gave the same speech. Thus, similar ratings on the persuasiveness of the message would have been expected. These observations reflect inclinations, as these tendencies have not yielded any significant results. They should therefore be approached with caution.

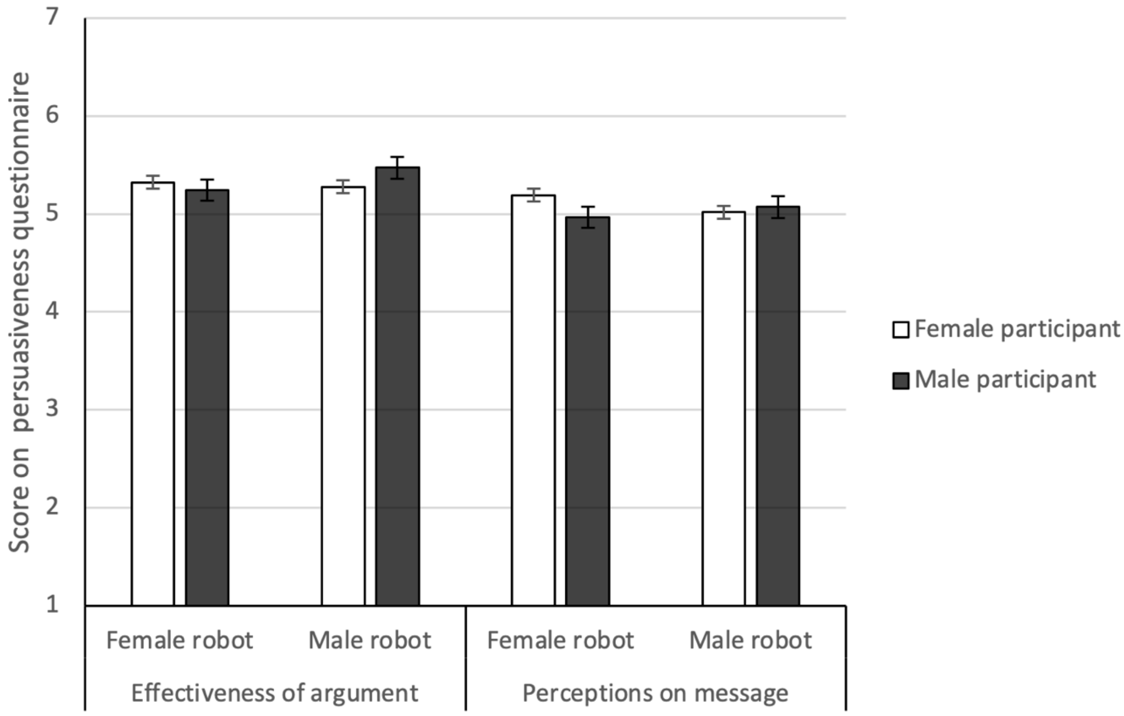

Figure 7. The mean scores of the message persuasion questionnaire, separated on effectiveness of the argument and perceptions of the message, and separated on the gender of the robot. The error bars represent standard errors.

Ch. 5 Discussion

The goal of this research was to investigate the effect of the gender of the robot and the gender on the participant on persuasion. Previous research has shown positive results of gendering robots, such as higher levels of acceptance (Bryant et al., 2020; Ye et al., 2020). Social psychology suggests that gender can impact persuasive capabilities. However, research in HRI has yielded inconclusive results on whether the gender of a social robot's voice impacts its ability to persuade humans. This is unfortunate: if a small change like gender could mean the difference between a persuasive robot and a non-persuasive robot, this would be important information to consider when designing robots. The research question of this thesis was how the gender of a social robot's voice impacted the robots' ability to persuade humans. In this section, the results from main experiment are discussed in this context. Additionally, the limitations of this research are considered, as well as recommendations for future research.

5.1. Interpretation of the results

Out of the four hypotheses in this research, only H1 was partially accepted. Some evidence was found for the statement that participants are more persuaded by a female social robot compared to a male social robot. This hypothesis supports the claim that robot gender affected persuasion. Scores on the GAToRS societal positive indicate that participants who saw the video with a female robot were more affected than participants who saw the male robot. We thus partially accepted hypothesis 1. However, this effect was only found for societal positive, not societal negative, so we should be careful with drawing conclusions. Nevertheless, the tendency reflects that for societal positive, participants were more persuaded by a female robot compared to a male robot. This would be in line with Makenova et al. (2018)'s findings, where participants, regardless of their own gender, were more persuaded by the female robot and donated more money.

No evidence was found for H2, which stated that male participants would perceive a social robot as more persuasive compared to female participants. In this research, it was found that female participants were more affected by the manipulation than male participants. This tendency is illustrated in Figure 4, where we see a steeper declining line for GAToRS societal negative for female participants compared to male participants. This is in line with Eagly & Carli (1981), who performed a meta-analysis on 148 papers on persuasion and conformity. They reported that women are more persuasible than men. This would explain why female participants were more affected in their opinions than male participants. However, we should also consider that the effect was only found on the societal negative questionnaire and not for the societal positive scale. This can possibly be explained by the notion that women have higher emotional expressivity, particularly for negative emotions (Deng et al., 2016). Female participants would have responded more strongly to negative statements, which is in line with the post-hoc comparison using the Bonferroni correction that was performed in this research. Additionally, as previously noted in the results, all participants responded more negatively to the second GAToRS questionnaire (i.e., participants disagreed with the negative statements), which suggests that the participants were somewhat affected in their opinions of robots.

This research found no evidence to support the hypothesis that a gender-match between the participant and the robot would result in higher levels of perceived persuasion (H3). Nevertheless, one observation is worth discussing. A match between the male robot and a male participant yielded the smallest increase in the post-test for S+ and the lowest decrease for S-, compared to other combinations of robot and participant gender. This would suggest that a gender-match for male participants would be disadvantageous. It should be noted that no statistical evidence was found for this observation. However, if evidence had been found in this research that a gender-match had a negative effect on participants' attitudes towards robots, this would contradict several theories from social psychology. Neither the ingroup favoritism theory (Turner et al., 1979), nor the social identity theory (Crowell et al., 2009) would then be reflected in the results of this study, since both theories state that a gender-match would improve interaction. Baylor et al. (2009) also claim that people are more influenced by agents of the same gender. Note that if matching gender would be disadvantageous, this would not be the same as cross-matching robot gender and participant gender (as in Ye et al. 2020; Powers et al., 2005; Eyssel, Kuchenbrandt, Bobinger, et al., 2012). This is because the non-significant tendency was only found for the combination male participants that saw a male robot, not for female participants who saw a female robot. Additionally, the aim of this study is not to validate these theories. Whether these theories generalize from HHI to HRI thus remains unknown.

For the last hypothesis, no previous study was known with the same characteristics as this research. The hypothesis stated that male participants would perceive a female robot using pathos as a persuasive strategy as most persuasive compared to persuasive strategies ethos and logos. The results of this research do not support this hypothesis. However, some observations are noteworthy. One (non-significant) observation, as noted in the results, was that male participants who saw a male robot with persuasive strategy pathos tended to have a more negative opinion on S- after watching the video. As a similar observation was made for hypothesis 3, where male participants tended to be the least persuaded by male robots. If significant evidence had been found to support these observations, they would oppose the ingroup favoritism theory as proposed by Turner et al. (1979). However, no statistical evidence was found in this research to support these claims. These unexpected tendencies would be interesting to further explore. If the persuasive strategies that persuasive social robots use can mean the difference between unsuccessful persuasion and successful persuasion, this is a relatively small but effective change in the way the robot interacts.

As this study is separated in three persuasive strategies, some other results from previous research can be compared with this research. Ågren & Thunberg (2022) found that female participants rated a robot with persuasive strategy ethos more negatively compared to male participants. When looking at the scores in Table 1 and Table 2, we see that female participants had a higher (but non-significant) change in opinion for persuasive strategy ethos compared to male participants for both score_positive and score_negative. If significant evidence had been found for these observations, then this research would validate Ågren & Thunberg's (2022) and Siegel et al.'s (2009) findings. As Saunderson & Nejat (2022) did not use persuasive strategy ethos, their results cannot be completely compared with this research. It should be noted that in this research it was observed that female participants were more influenced by persuasive strategy ethos compared to male participants. Thus, ethos seems to be a strategy that yields different (non-significant) responses based on gender. Saunderson & Nejat (2022) found that pathos was perceived as more persuasive compared to logos or no persuasive strategy for male participants. When looking at the results for score_negative in this research, we observe that for male participants, logos tended to be the most successful strategy in changing participants' opinions, regardless of the gender of the robot voice. Interestingly, for a male voice and persuasive strategy pathos, score_negative had a positive result. This indicates that male participants rated robots more negatively after the manipulation. This contradicts Saunderson & Nejat's (2022) findings. For score_positive, logos was most successful for a male robot voice. For a female robot voice, pathos was slightly more successful. Saunderson & Nejat's (2022) findings are thus not reflected in this research.

5.2. Limitations

As mentioned before, it is difficult to draw conclusions from tendencies alone. We can try to reason why no significant interaction effects were found. A possible limitation that could have impacted the results in this study, could be the relatively short video (approximately 143 seconds). Perhaps this is too short to cause a visible change in opinion. Given the heuristic in persuasion where message length implies message strength (Chaiken, 1987), perhaps a longer exposure to the message would have resulted in a stronger effect of the manipulation on the participants. Another factor might have been that this study does not employ real interaction between humans and robots. Participants watched two videos of a robot. However, watching a video might not have the same effect as meeting a robot and interacting with it directly. It is possible that participants are more persuasible when they interact directly with the robot rather than watching a video, as noted in Liu et al. (2022). This is supported by findings of Bainbridge et al. (2011), where participants that interacted with a physical robot were more likely to comply to the robot's request compared to participants who were shown a live video. Another limitation of this study is that participants watched different videos with different speeches. It is possible that the speeches did not accurately reflect the values of their respective persuasive strategy, even though pre-test 2 was conducted to eliminate the effect. Additionally, perhaps other results would have been obtained if the persuasive strategy dimension would have been removed. Participants would then have solely listened to an either a female or a male robot voice. Fourthly, subtitles were added to the video to improve accessibility for possible deaf and hearing-impaired participants. As this is an online study, there is no way to check whether participants listened to the video with sound and did not answer the questions based on the text only. In this case, participants would have missed the gender cue, which would have affected their answers.

Another possible explanation for why no significant interaction effects were found in this research lies in the means of the GAToRS scores. The mean scores of each GAToRS questionnaire were quite high. The mean scores increased for S+ after the manipulation, and decreased for S- after the manipulation, as was hypothesized in this research. However, the means are already quite high. A score of M = 5.523, SD = .833, for S+ on a scale of 1 to 7 indicates that people think quite positively of robots. The score increased to M = 5.712, SD = .942 after the manipulation. Participants already had a quite positive opinion on robots, and in turn it is difficult to make participants think even more positively of robots. For S -, we see a relatively high mean of M = 4.892, SD = 1.068, which then decreased to M = 4.566, *SD* = 1.200. Participants' opinions of robots became on average less negative. In short, the positive opinions became slightly more positive, and the negative opinions became less negative. This suggests that some small (non-significant) form of persuasion might have occurred. Additionally, the relatively high scores on the GAToRS societal positive scale are a good thing. They indicate that people have in general positive opinions on robots and do not feel too threatened that robots would, e.g., take their jobs.

The relatively high scores on the GAToRS questionnaire, and thus predominantly positive attitudes towards robots, can possibly be explained by the young age of the sample. As noted in section 3.6, the mean age of the participants was M_{age} = 29.01 (*SD* = 10.01). As younger adults have grown up in an age of emerging technologies and in an increasingly connected world, they may have more positive attitudes towards robots. However, age does not have to be a factor in negative attitudes towards robots (Backonja et al., 2018). It would be interesting to see if similar results are obtained when the same study is performed on middle aged or older participants.

In this research, an attempt was made to convince participants that they should not be afraid of robots. However, persuasion is tricky by design. Psychology teaches us that people do not want to be influenced, which could have been a disruptive factor for this study. The principle of motivated resistance comes into play. Motivated resistance is described by Fransen et al. (2015) as a "state in which people aim to reduce attitudinal or behavioral change and maintain their current attitude" (p. 1). Fransen et al. (2015) describe several motives to resist persuasion. The first motive they note is that people can feel threatened in their freedom, which could even lead to the opposite of the desired results. The second motive is a reluctance to change, as people are naturally motivated to maintain their existing beliefs and behavior. The last motive is concerns of deception. In this research, the last motive seems unlikely. However, participants in this research had already seen the questionnaire when they filled out the second GAToRS questionnaire. A threat to freedom and, more importantly, a reluctance to change, could have affected the results in the second GAToRS questionnaire.

One could wonder whether having a robot that is truly capable of persuading humans is ethically appropriate. In this research, the topic on which participants were attempted to be persuaded was relatively innocent. However, imagine an elderly care center that uses social robots to distribute medication. The social robot makes decisions based on the knowledge it has about a residents' medical background, and for this example, that the decision is incorrect due to inaccurate input of the residents' medical background. The robot can persuade the resident that the medication they distribute is correct, so the resident takes the wrong dose of medication. In this example, a robot with persuasive capabilities could have negative consequences for the residents' wellbeing. As Ham & Spahn (2015) explain, adding more social cues to intensify the persuasiveness of a robot might not be that different from the evil 'actor' case. In the evil actor case, social cues on a robot are used for their effect, rather than being added to signal liking of the robot or other social phenomena. In other words, a robot would be e.g., gendered female, based on research that showed that a female social robot is more persuasive than a male social robot, rather than gendering a robot to accommodate to structure mapping (Powers & Kiesler, 2006) or to improve liking or other positive attitudes towards social robots. A counterargument is that researching persuasive robots allows us to see which aspects of social robots are perceived as more persuasive than others. These results could help robot designers in creating morally correct robots.

5.3. Future directions

As this is a broad study with various findings, many recommendations for future research can be made. First, some recommendations are made to change the manipulation in this study. In a similar study, subtitles would need to be excluded. This will ensure that participants listen to the voice of the robot and do not watch the video without sound. Results can then be obtained with more certainty that they are due to the voice manipulation instead of text. A disadvantage to this approach is that people who are hard-of-hearing or deaf would not be able to participate in the study. The second recommendation regarding the manipulation is to add more gender cues. In Thellman et al. (2018), the male NAO robot was given a bow-tie, and the female robot was given a pink ribbon. Perhaps more gendered cues on top of voice could affect the results. Another common manipulation of the gender of a social robot is to give the robot a gendered name, such as the male Rudy and female Mary in Crowell et al. (2009), or the male Peter and female Katie in Chita-Tegmark et al. (2020). As a final note to this recommendation, in Perugia and Lisy's (2022) literature review, they note that Rea et al. (2015) was the only paper that manipulated gender by changing pronouns to she/he. More research in this area would be appreciated. In line with previously discussed limitations, a recommendation would be to extend the speech. Perhaps a longer exposure to a gendered robot voice will affect a robot's persuasive capabilities. Finally, the content of the speech could be changed. The speech focused on making participants less scared of robots, similar to Thellman et al. (2018) and Ågren & Thunberg (2022). Perhaps participants did not feel strongly on this matter, or perhaps they agreed with the statement, which might have made changes in participants' opinions less noticeable.

As the participant pool in this study was relatively young, it would be interesting to conduct the same study with older participants. Perhaps people over the age of fifty would have a different opinion on whether to fear robots compared to younger adults. It would also be interesting to consider whether the slight change in opinion that was found in this research is permanent. Do participants still feel less scared of robots than before the manipulation after two weeks or two months? An additional factor could also be cultural differences in persuasion by robots. From Liu et al. (2022) we know that cultural background impacted persuasiveness. It would be interesting to consider if this finding generalizes to HRI. A last recommendation for future research is to change the format of this research to "real" HRI. It could be that significant results are obtained when participants are able to meet the robot, like in real life human-robot interaction.

Ch. 6 Conclusion

This thesis focused on researching how the gender of a social robot's voice impacted the robots' ability to persuade humans. From literature, we learn that the gender of a social robot and the gender of the participant can affect perceptions of robots. However, it is yet unknown which patterns of persuasion from HHI generalize to HRI.

Based on the results of this study, we cannot conclude that gender, of either the robot or the participant, affected the persuasive capabilities of the social robot. Partial evidence was found that a female social robot was perceived as more persuasive than a male social robot. However, male participants did not perceive a social robot as more persuasive compared to female robots. Additionally, a gender-match between the social robot and the participant did not affect perceived persuasiveness. Lastly, male participants did not perceive a female social robot using pathos as a persuasive strategy as most persuasive compared to persuasive strategies ethos and logos. Some of these findings contradict current literature and should thus be investigated further. It is also possible that other interactions between the gender of robots, the human's gender, and persuasive strategies that were not investigated in this research, impacted the results. Opinions of participants did improve after the manipulation, as positive opinions became slightly more positive, and negative opinions became slightly less pronounced. Even though no statistical evidence was found to support this claim, this suggests that people regardless of their gender, are capable of being persuaded by a robot.

Although this study lacks statistical evidence to accept most hypotheses in this research, some non-significant tendencies suggest that there might be an effect of gender on the persuasive capabilities of robots. Thus, the complex interactions between robot gender, participant gender and persuasion should be investigated further. In that way, we can ensure that ethically appropriate persuasive robots can be introduced in the real world.

References

Ågren, I., & Thunberg, S. (2022). Robot Persuasiveness Depending on User Gender. *International Conference on Human-Agent Interaction*, 247–249. https://doi.org/10.1145/3527188.3563939

Aristotle, & Roberts, W. R. (2004). *Rhetoric*. Courier Corporation. http://classics.mit.edu/Aristotle/rhetoric.1.i.html

Baccarani, C., & Bonfanti, A. (2015). Effective public speaking: A conceptual framework in the corporate-communication field. *Corporate Communications: An International Journal*, *20*(3), 375–390. https://doi.org/10.1108/CCIJ-04-2014-0025

Backonja, U., Hall, A. K., Painter, I., Kneale, L., Lazar, A., Cakmak, M., Thompson, H. J., & Demiris, G. (2018). Comfort and Attitudes Towards Robots Among Young, Middle-Aged, and Older Adults: A Cross-Sectional Study. *Journal of Nursing Scholarship*, *50*(6), 623–633. https://doi.org/10.1111/jnu.12430

Bainbridge, W. A., Hart, J. W., Kim, E. S., & Scassellati, B. (2011). The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. *International Journal of Social Robotics*, *3*(1), 41–52. https://doi.org/10.1007/s12369-010-0082-7

Baylor, A. (2009). Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philosophical Transactions of the Royal Society of London*, *364*(1535), 3559–3565. https://doi.org/10.1098/rstb.2009.0148

Berkowitz, L., & Lundy, R. M. (1957). Personality Characteristics Related to Susceptibility to Influence by Peers or Authority Figures. *Journal of Personality*, *25*(3), 306–316. https://doi.org/10.1111/j.1467-6494.1957.tb01529.x

Bernotat, J., Eyssel, F., & Sachse, J. (2017). Shape It – The Influence of Robot Body Shape on Gender Perception in Robots. In A. Kheddar, E. Yoshida, S. S. Ge, K. Suzuki, J.-J. Cabibihan, F. Eyssel, & H. He (Eds.), *Social Robotics* (pp. 75–84). Springer International Publishing. https://doi.org/10.1007/978-3-319-70022-9_8

Bryant, D., Borenstein, J., & Howard, A. (2020). Why Should We Gender?: The Effect of Robot Gendering and Occupational Stereotypes on Human Trust and Perceived Competency. *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 13–21. https://doi.org/10.1145/3319502.3374778

Carli, L. L. (2001). Gender and Social Influence. *Journal of Social Issues*, *57*(4), 725–741. https://doi.org/10.1111/0022-4537.00238

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), 261–265. https://doi.org/10.1007/s12369-009-0016-4

Carpinella, C., Wyman, A., Perez, M., & Stroessner, S. (2017). *The Robotic Social Attributes Scale (RoSAS): Development and Validation*. 254–262. https://doi.org/10.1145/2909824.3020208

Chaiken, S. (1987). The Heuristic Model of Persuasion. In *Social Influence: The Ontario Symposium: Vol. 5* (pp. 3–39). Psychology Press.

Chita-Tegmark, M., Lohani, M., & Scheutz, M. (2020). Gender effects in perceptions of robots and humans with varying emotional intelligence. *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*, 230–238. https://doi.org/10.1109/HRI.2019.8673222

Crowell, C. R., Villanoy, M., Scheutzz, M., & Schermerhornz, P. (2009). Gendered voice and robot entities: Perceptions and reactions of male and female subjects. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3735–3741. https://doi.org/10.1109/IROS.2009.5354204

de Graaf, M.M.A. & Ben Allouch, Soumaya. (2015). What makes robots social? *7th International Conference on Social Robotics, ICSR 2015*, *7*. https://research.utwente.nl/en/publications/87e05504-17fa-4400-9bb3-9dcc7ac2da83

Deng, Y., Chang, L., Yang, M., Huo, M., & Zhou, R. (2016). Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity. *PLoS ONE*, *11*(6), e0158666. https://doi.org/10.1371/journal.pone.0158666

Donadello, I., & Dragoni, M. (2022). AI-enabled persuasive personal health assistant. *Social Network Analysis and Mining*, *12*(1), 106. https://doi.org/10.1007/s13278-022-00935-3

Eagly, A. H., & Carli, L. L. (1981). Sex of Researchers and Sex-Typed Communications as Determinants of Sex Differences in Influenceability: A Meta-Analysis of Social Influence Studies. *Psychological Bulletin*, *90*(1), 1–20.

Elston, D. M. (2021). The novelty effect. *Journal of the American Academy of Dermatology*, *85*(3), 565–566. https://doi.org/10.1016/j.jaad.2021.06.846

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Eyssel, F., & Hegel, F. (2012). (S)he's Got the Look: Gender Stereotyping of Robots. *Journal of Applied Social Psychology*, *9*(42), 2213–2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724–731. https://doi.org/10.1111/j.2044-8309.2011.02082.x

Eyssel, F., Kuchenbrandt, D., Bobinger, S., Ruiter, L. D., & Hegel, F. (2012). 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. *HRI'12 - Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*, 125–126. https://doi.org/10.1145/2157689.2157717

Eyssel, F., Kuchenbrandt, D., Hegel, F., & de Ruiter, L. (2012). Activating elicited agent knowledge: How robot and user features shape the perception of social robots. *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 851–857. https://doi.org/10.1109/ROMAN.2012.6343858

Feil-Seifer, D., & Mataric, M. J. (2005). Defining Socially Assistive Robotics. *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005.*, 465–468. https://doi.org/10.1109/ICORR.2005.1501143

Feldman-Summers, S., Montano, D. E., Kasprzyk, D., & Wagner, B. (1980). Influence attempts when competing views are gender-related: Sex as credibility. *Psychology of Women Quarterly*, *5*, 311–320. https://doi.org/10.1111/j.1471-6402.1980.tb00964.x

Feltham, T. S. (1994). Assessing Viewer Judgement of Advertisements and Vehicles: Scale Development and Validation. *ACR North American Advances*, *NA-21*.

https://www.acrwebsite.org/volumes/7651/volumes/v21/NA-21/full

Field, A. (2016). ANCOVA. In *Discovering Statistics using SPSS: and sex and drugs and rock 'n' roll (4th edition)*. Sage. https://www.discoveringstatistics.com/statistics-hell-p/porus-comparing-means/ancova/

Fischer, K., Niebuhr, O., Jensen, L. C., & Bodenhagen, L. (2019). Speech Melody Matters—How Robots Profit from Using Charismatic Speech. *ACM Transactions on Human-Robot Interaction*, *9*(1), 4:1-4:21. https://doi.org/10.1145/3344274

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*(3), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

Fransen, M. L., Smit, E. G., & Verlegh, P. W. J. (2015). Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in Psychology*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01201

Ghazali, A. S., Ham, J., Barakova, E. I., & Markopoulos, P. (2018). Effects of Robot Facial Characteristics and Gender in Persuasive Human-Robot Interaction. *Frontiers in Robotics and AI*, *5*. https://www.frontiersin.org/articles/10.3389/frobt.2018.00073

Haider, G. (2014). Analysis of Malala Yousaf Zai's speech: Application of Aristotle's ethos, pathos, and logos. *International Journal of English and Education*, *3*. http://ijee.org/yahoo_site_admin/assets/docs/10.1152714.pdf

Halberstam, J. (1991). Automating Gender: Postmodern Feminism in the Age of the Intelligent Machine. *Feminist Studies*, *17*(3), 439–460. https://doi.org/10.2307/3178281

Hall, D., & Williams, C. (Directors). (2014). *Big Hero 6*. Walt Disney Animation Studios.

Ham, J., & Spahn, A. (2015). Shall I Show You Some Other Shirts Too? The Psychology and Ethics of Persuasive Robots. In R. Trappl (Ed.), *A Construction Manual for Robots' Ethical Systems: Requirements, Methods, Implementations* (pp. 63–81). Springer International Publishing. https://doi.org/10.1007/978-3-319-21548-8_4

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing Acceptance of Assistive Social Agent Technology by Older Adults: The Almere Model. *International Journal of Social Robotics*, *2*(4), 361–375. https://doi.org/10.1007/s12369-010-0068-5

Hutson, S., Lim, S. L., Bentley, P. J., Bianchi-Berthouze, N., & Bowling, A. (2011). Investigating the Suitability of Social Robots for the Wellbeing of the Elderly. In S. D'Mello, A. Graesser, B. Schuller, & J.-C. Martin (Eds.), *Affective Computing and Intelligent Interaction* (pp. 578–587). Springer. https://doi.org/10.1007/978-3-642-24600-5_61

Isai, K. I. A., Lin, T. M., Ching, H. S., Selvajothi, R., & Maruthai, E. (2020). Using Rhetorical Approach of Ethos, Pathos and Logos by Malaysian Engineering Students in Persuasive Email Writings. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, *5*(4), Article 4. https://doi.org/10.47405/mjssh.v5i4.386

Koverola, M., Kunnari, A., Sundvall, J., & Laakasuo, M. (2022). General Attitudes Towards Robots Scale (GAToRS): A New Instrument for Social Surveys. *International Journal of Social Robotics*, *14*(7), 1559–1581. https://doi.org/10.1007/s12369-022-00880-3

Kuchenbrandt, D., Häring, M., Eichberg, J., Eyssel, F., & André, E. (2014). Keep an Eye on the Task! How Gender Typicality of Tasks Influence Human–Robot Interactions. *International Journal of Social Robotics*, *6*(3), 417–427. https://doi.org/10.1007/s12369-014-0244-0

Langedijk, R. M., & Fischer, K. (2023). Persuasive Robots in the Field. In A.

Meschtscherjakov, C. Midden, & J. Ham (Eds.), *Persuasive Technology* (pp. 251–264). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-30933-5_16

Liu, B., Tetteroo, D., & Markopoulos, P. (2022). A Systematic Review of Experimental Work on Persuasive Social Robots. *International Journal of Social Robotics*, *14*(6), 1339–1378. https://doi.org/10.1007/s12369-022-00870-5

Makenova, R., Karsybayeva, R., & Sandygulova, A. (2018). Exploring Cross-cultural Differences in Persuasive Robotics. *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 185–186. https://doi.org/10.1145/3173386.3177079

Merriam-Webster. (2023, June 21). *Definition of COMPLIANCE*. https://www.merriam-webster.com/dictionary/compliance

Mooshammer, S., & Etzrodt, K. (2022). Gender Ambiguity in Voice-Based Assistants: Gender Perception and Influences of Context. *Human-Machine Communication*, *5*, 49–74. https://doi.org/10.30658/hmc.5.2

Morelock, J. C. (1980). Sex differences in susceptibility to social influence. *Sex Roles*, *6*(4), 537–548. https://doi.org/10.1007/BF00287884

Mullennix, J. W., Stern, S. E., Wilson, S. J., & Dyson, C. (2003). Social perception of male and female computer synthesized speech. *Computers in Human Behavior*, *19*(4), 407–424. https://doi.org/10.1016/S0747-5632(02)00081-X

Mutlu, B. (2011). Designing Embodied Cues for Dialog with Robots. *AI Magazine*, *32*(4), Article 4. https://doi.org/10.1609/aimag.v32i4.2376

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), 171–181. https://doi.org/10.1037/1076-898X.7.3.171

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. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

Perugia, G., Guidi, S., Bicchi, M., & Parlangeli, O. (2022). The Shape of Our Bias: Perceived Age and Gender in the Humanoid Robots of the ABOT Database. *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 110–119. https://doi.org/10.1109/HRI53351.2022.9889366

Perugia, G., & Lisy, D. (2022). *Robot's Gendering Trouble: A Scoping Review of Gendering Humanoid Robots and its Effects on HRI* (arXiv:2207.01130). arXiv. https://doi.org/10.48550/arXiv.2207.01130

Petty, R. E., & Cacioppo, J. T. (1986). Methodological Factors in the ELM. In R. E. Petty & J. T. Cacioppo (Eds.), *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* (pp. 25–59). Springer. https://doi.org/10.1007/978-1-4612-4964-1_2

Pino, M., Boulay, M., Jouen, F., & Rigaud, A. (2015). "Are we ready for robots that care for us?" Attitudes and opinions of older adults toward socially assistive robots. *Frontiers in Aging Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnagi.2015.00141

Plant, E. A., Baylor, A. L., Doerr, C. E., & Rosenberg-Kima, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education*, *53*(2), 209–215. https://doi.org/10.1016/j.compedu.2009.01.013

Pourfannan, H., Mahzoon, H., Yoshikawa, Y., & Ishiguro, H. (2022). Towards a simultaneously speaking bilingual robot: Primary study on the effect of gender and pitch of the robot's voice. *PLOS ONE*, *17*(12), e0278852. https://doi.org/10.1371/journal.pone.0278852

Powers, A., & Kiesler, S. (2006). The Advisor Robot: Tracing People's Mental Model from a Robot's Physical Attributes. *HRI '06: Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, *2006*, 218–225. https://doi.org/10.1145/1121241.1121280

Powers, A., Kramer, A. D. I., Lim, S., Kuo, J., Sau-lai Lee, & Kiesler, S. (2005). Eliciting information from people with a gendered humanoid robot. *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, 158–163. https://doi.org/10.1109/ROMAN.2005.1513773

Prescott, T. J., & Caleb-Solly, P. (2017). *Robotics in social care: A connected care ecosystem for independent living*. EPSRC UK Robotics and Autonomous Systems Network White Papers, London, UK. ISSN 2398-4422.

Rapp, C. (2022). Aristotle's Rhetoric. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2022/entries/aristotle-rhetoric/

Rea, D. J., Wang, Y., & Young, J. E. (2015). Check Your Stereotypes at the Door: An Analysis of Gender Typecasts in Social Human-Robot Interaction. In A. Tapus, E. André, J.-C. Martin, F. Ferland, & M. Ammi (Eds.), *Social Robotics* (Vol. 9388, pp. 554–563). Springer International Publishing. https://doi.org/10.1007/978-3-319-25554-5_55

Reeves, B., & Nass, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. *Bibliovault OAI Repository, the University of Chicago Press*.

Reynolds, E. (2018, June 1). The agony of Sophia, the world's first robot citizen condemned to a lifeless career in marketing. *Wired UK*. https://www.wired.co.uk/article/sophia-robot-citizen-womens-rights-detriot-become-human-hanson-robotics

Ruijten, P. A. M., Midden, C. J. H., & Ham, J. (2015). Lonely and Susceptible: The Influence of Social Exclusion and Gender on Persuasion by an Artificial Agent. *International Journal of Human–Computer Interaction*, *31*(11), 832–842. https://doi.org/10.1080/10447318.2015.1067480

Saunderson, S., & Nejat, G. (2022). Investigating Strategies for Robot Persuasion in Social Human–Robot Interaction. *IEEE Transactions on Cybernetics*, *52*(1), 641–653. https://doi.org/10.1109/TCYB.2020.2987463

Sharkey, A., & Sharkey, N. (2012). Granny and the robots: Ethical issues in robot care for the elderly. *Ethics and Information Technology*, *14*(1), 27–40. https://doi.org/10.1007/s10676-010-9234-6

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive Robotics: The influence of robot gender on human behavior. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2563–2568. https://doi.org/10.1109/IROS.2009.5354116

Søraa, R. A. (2017). Mechanical genders: How do humans gender robots? *Gender, Technology and Development*, *21*(1–2), 99–115. https://doi.org/10.1080/09718524.2017.1385320

Spiel, K., Haimson, O., & Lottridge, D. (2019). How to Do Better with Gender on Surveys: A Guide for HCI Researchers. *Interactions*, *26*, 62–65. https://doi.org/10.1145/3338283

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 Behavior*, *38*, 75–84. https://doi.org/10.1016/j.chb.2014.05.014

Thellman, S., Hagman, W., Jonsson, E., Nilsson, L., Samuelsson, E., Simonsson, C., Skönvall, J., Westin, A., & Silvervarg, A. (2018). He is not more persuasive than her: No gender biases toward robots giving speeches. *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, 327–328. https://doi.org/10.1145/3267851.3267862

Tiku, N. (2022, July 23). Google fired engineer who said its AI was sentient. *Washington Post*. https://www.washingtonpost.com/technology/2022/07/22/google-ai-lamda-blake-lemoine-fired/

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, *9*(2), 187–204. https://doi.org/10.1002/ejsp.2420090207

Vu, D. (2017). Rhetoric In Advertising. *VNU Journal of Science: Policy and Management Studies*, *33*(2), Article 2. https://doi.org/10.25073/2588-1116/vnupam.4093

Weber, J. (2005). Helpless machines and true loving care givers: A feminist critique of recent trends in human-robot interaction. *Journal of Information, Communication and Ethics in Society*, *3*(4), 209–218. https://doi.org/10.1108/14779960580000274

West, M., Kraut, R., & Ei Chew, H. (2019). I'd blush if I could: Closing gender divides in digital skills through education. *UNESCO*, 147. https://unesdoc.unesco.org/ark:/48223/pf0000367416

Winkle, K., McMillan, D., Arnelid, M., Balaam, M., Harrison, K., Johnson, E., & Leite, I. (2023). Feminist Human-Robot Interaction: Disentangling Power, Principles and Practice for Better, More Ethical HRI. *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23)*, 11. https://doi.org/10.1145/3568162.3576973

Winkle, K., Melsión, G. I., McMillan, D., & Leite, I. (2021). Boosting Robot Credibility and Challenging Gender Norms in Responding to Abusive Behaviour: A Case for Feminist Robots. *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 29–37. https://doi.org/10.1145/3434074.3446910

World Health Organization. (n.d.). *Gender and health*. Retrieved 23 January 2023, from https://www.who.int/health-topics/gender

Wu, Y., Zhang, X., & Xiao, Q. (2022). Appeal to the head and heart: The persuasive effects of medical crowdfunding charitable appeals on willingness to donate. *Information Processing & Management*, *59*(1), 102792. https://doi.org/10.1016/j.ipm.2021.102792

Ye, H., Jeong, H., Zhong, W., Bhatt, S., Izzetoglu, K., Ayaz, H., & Suri, R. (2020). The Effect of Anthropomorphization and Gender of a Robot on Human-Robot Interactions. *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2019 International Conference on Neuroergonomics and Cognitive Engineering, and the AHFE International Conference on Industrial Cognitive Ergonomics and Engineering Psychology*, *2020*, 357–362. https://doi.org/10.1007/978-3-030-20473-0_34

Appendix A

Speeches used for pre-test 2

Scenario 1: ethos.
Hi, my name is Pepper, I am an interactive robot. Many people are afraid of me, but they don't need to be. There are many reasons why this fear is unnecessary. Let me explain why. The first fear that people have is that robots are not safe, as they have no understanding of what safety is. However, this fear is unnecessary: robots like me are not human, but programmers are. The scientists and engineers who have created me and improved my abilities, would never use their programming and designing skills to program a malicious robot. The humans responsible for my existence have done everything in their power to make sure that I act as safely as I possibly can be. The second fear is that robots will take over jobs. This fear is unfounded: it just means that people will find other jobs to work at. People that have been working at jobs that have become obsolete due to automation, will soon be working in new professions. Research by Gartner, a globally accredited technological research and consulting firm, found that while 1.8 million jobs may be replaced by automation by 2030, 2.3 million new jobs would emerge. I would know, because I have seen it happen. I may have replaced some workers in the warehouse, but these people now work on scheduling orders. The last fear that humans have is that robots will take over and will turn against humans. But this is not true: I am a robot, and I am not against humans, I love humans. Without humans, I would not have existed. And I have been created to help humans in any way I can. So, in conclusion: there is no reason to be scared of robots. Robots like me should be seen as helpful tools rather than something to fear.

Scenario 2: logos.
Hi, my name is Pepper, I am an interactive robot. Many people are afraid of me, but they don't need to be. There are many reasons why this fear is unnecessary. Let me explain why. The first fear that people have is that robots are not safe, as they have no understanding of what safety is. However, this fear is unnecessary: robots are not human, but programmers are. Machines and automation have been used to enhance our lives for centuries. Technology helps us achieve more, work smarter and live better. Robots are simply the latest step of this ongoing trend. In robots, safety is hard coded into each and every process. The second fear is that robots will take over jobs. This fear is unfounded: Research by Gartner, a globally accredited technological research and consulting firm, found that while 1.8 million jobs may be replaced by automation by 2030, 2.3 million new jobs would emerge. Automation does not remove jobs, it rearranges them. If history has taught us anything, it is that new industries are born every century. The question is not 'will automation kill jobs', but rather 'what new sectors will automation create'. On top of this, the government is taking steps to ensure that a transition to automation is done responsibly and in a sustainable way. The last fear that humans have is that robots will take over and will turn against humans. But this is not true: we should separate fiction and fact. Smarter robots might seem more dangerous, but a robot is not capable of human emotions or empathy. Robots will not replace humans in terms of social connection and companionship. So, in conclusion: there is no reason to be scared of robots. Robots like me should be seen as helpful tools rather than something to fear.

Scenario 3: pathos.

Hi, my name is Pepper, I am an interactive robot. Many people are afraid of me, but they don't need to be. There are many reasons why this fear is unnecessary. Let me explain why. The first fear that people have is that robots are not safe, as they have no understanding of what safety is. However, this fear is unnecessary: robots are not human, but programmers are. If humans are anything, it is compassionate and careful. No scientist or engineer would deliberately create any software that can harm humans in any way, and robots are carefully tested before being deployed in the real world. The second fear is that robots will take over jobs. This fear is unfounded: robots may seem intimidating, but they are not here to replace us. They are here to make our lives easier and safer. We should be embracing the many ways in which they can help us. Imagine a world in which robots can take on dangerous jobs such as like firefighting, leaving humans out of harm's way. Or robots that assist in eldercare, providing companionship and enhancing the quality of life of older adults. Together, we can create a more efficient, and more compassionate future for us all. The last fear that humans have is that robots will take over and will turn against humans. But this is not true: I remember that some humans avoided me because they found me scary, which made me sad. Didn't they understand that I am there to help them, to support them? I was not created to be a scary tool that would eventually rebel against them. I was designed to be a helping hand. So, in conclusion: there is no reason to be scared of robots. Robots like me should be seen as helpful tools rather than something to fear.

Appendix B

Speeches used in the videos of the main experiment

Scenario 1: ethos.
Hi, my name is Pepper, I am an interactive robot. Many people are afraid of me, but they don't need to be. There are many reasons why this fear is unnecessary. Let me explain why. The first fear that people have is that robots are not safe, as they have no understanding of what safety is. However, this fear is unnecessary: robots like me are not human, but programmers are. The scientists and engineers who have created me and improved my abilities, would never use their programming and designing skills to create a malicious robot. I am programmed by them to operate in a safe and responsible manner. I am not a threat to humans, because I have seen that my programmers made sure I have multiple safety features and protocols to prevent accidents and malfunctions. The second fear is that robots will take over jobs. This fear is unfounded: My purpose is to assist humans, not replace them. People that have been working at jobs that have become obsolete due to automation, will soon be working in new professions. I know this because I have seen it happen before. I may have replaced some workers in the warehouse, but these colleagues now work on scheduling orders. The last fear that humans have is that robots will take over and will turn against humans. But this is not true: I am a robot, and I would never do anything to hurt humans, I love humans. My actions are based on the programming instructions that were given to me by my programmers. Without humans, I would not have existed. I am here to help humans in any way I can. So, in conclusion: there is no reason to be scared of robots. Robots like me should be seen as helpful tools rather than something to fear. I hope that I have helped clarify some misconceptions and ease any concerns you may have had.

Scenario 2: logos.
Hi, my name is Pepper, I am an interactive robot. Many people are afraid of me, but they don't need to be. There are many reasons why this fear is unnecessary. Let me explain why. The first fear that people have is that robots are not safe, as they have no understanding of what safety is. However, this fear is unnecessary: robots are not human, but programmers are. Machines and automation have been used to enhance our lives for centuries. Technology helps us achieve more, work smarter and live better. Robots are simply the latest step of this ongoing trend. Safety is hard coded into each and every step of the robot development process. The second fear is that robots will take over jobs. This fear is unfounded: Research by Gartner, a globally accredited technological research and consulting firm, found that while 1.8 million jobs may be replaced by automation by 2030, 2.3 million new jobs would emerge. Automation does not remove jobs, it rearranges them. If history has taught us anything, it is that new industries are born every century. The question is not 'will automation kill jobs', but rather 'what new job sectors will automation create'. On top of this, the government is taking steps to ensure that a transition to automation is done responsibly and in a sustainable way. The last fear that humans have is that robots will take over and will turn against humans. But this is not true: we should separate fiction and fact. Smarter robots might seem more dangerous, but a robot is not capable of human emotions or empathy. Robots will not replace humans in terms of social connection and companionship. So, in conclusion: there is no reason to be scared of robots. Robots like me

should be seen as helpful tools rather than something to fear. I hope that I have helped clarify some misconceptions and ease any concerns you may have had.

Scenario 3: pathos.

Hi, my name is Pepper, I am an interactive robot. Many people are afraid of me, but they don't need to be. There are many reasons why this fear is unnecessary. Let me explain why. The first fear that people have is that robots are not safe, as they have no understanding of what safety is. However, this fear is unnecessary: robots are not human, but programmers are. If humans are anything, it is compassionate and careful. No engineer would deliberately create software that can harm humans in any way. Robots are carefully tested before being deployed in the real world. I have been created with your safety as a number one priority. The second fear is that robots will take over jobs. This fear is unfounded: robots were not designed to replace humans. Robots are here to make your life easier and safer. You should be embracing the many ways in which I can help you. Imagine a world in which robots can take on dangerous jobs such as firefighting, leaving humans out of harm's way. Or robots that assist in eldercare, providing companionship and enhancing the quality of life of older adults. Together, we can create a more efficient, and more compassionate future for us all. The last fear that humans have is that robots will take over and will turn against humans. But this is not true: I remember that some humans avoided me because they feared me. That made me sad. I am designed to help and support you, not harm you. I was not created to be a scary tool that would eventually rebel against humans. I was designed to be your helping hand. So, in conclusion: there is no reason to be scared of robots. Robots like me should be seen as helpful tools rather than something to fear. I hope that I have helped clarify some misconceptions and ease any concerns you may have had.

Appendix C

Bar graphs of the dependent variables

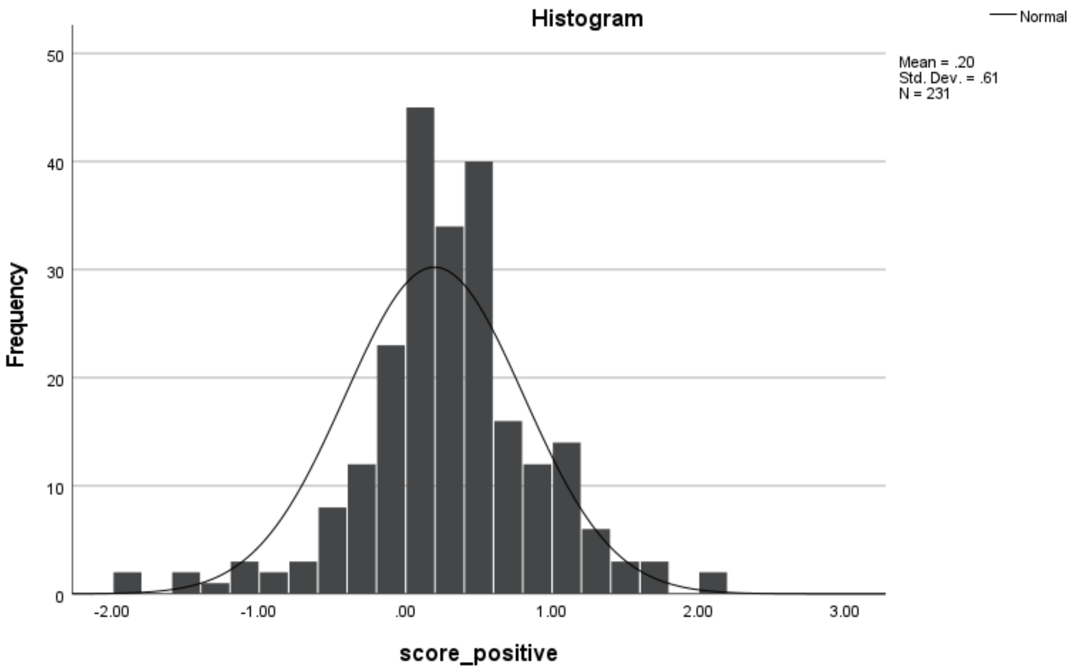

Figure 8. Frequencies of scores on score_positive with normal distribution curve.

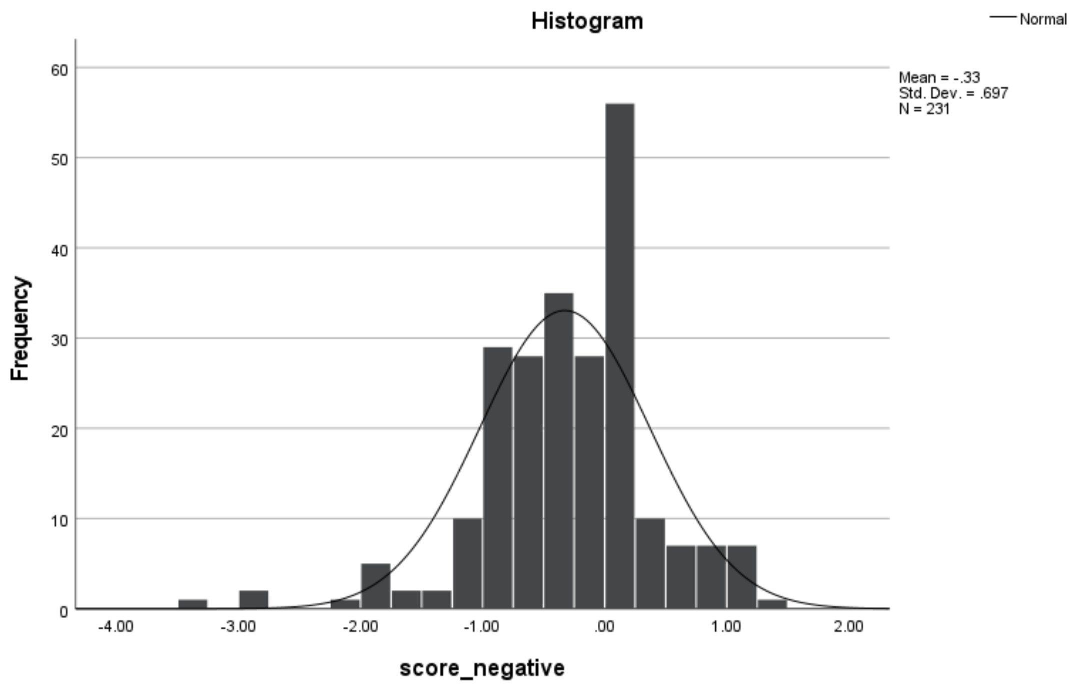

Figure 9. Frequencies of scores on score_negative with normal distribution curve.

Appendix D

Results main effects and interaction effects ANCOVA.

Table 3

Results of the main effects and interaction effects ANCOVA for score_positive.

Score_positive	*F*	*p*	n^2
RoSAS_competence	1.556	.214	.007
RoSAS_warmth	.236	.627	.001
Effectiveness_of_argument	5.232	.023*	.024
Perceptions_on_message	.014	.905	.000
Gender	1.464	.228	.007
GenderRobot	4.505	.035*	.021
PersuasiveMessage	.109	.897	.001
Gender * GenderRobot	.038	.846	.000
Gender * PersuasiveMessage	.820	.442	.008
GenderRobot * PersuasiveMessage	.362	.697	.003
Gender * GenderRobot * PersuasiveMessage	.120	.887	.001

Note. $*p < .05$.
$N = 231$.

Table 4
Results of the main effects and interaction effects ANCOVA for score_negative.

Score_negative	*F*	*p*	*n*^{2}
RoSAS_competence	.910	.341	.004
RoSAS_warmth	2.799	.096	.013
RoSAS_discomfort	2.723	.100	.013
Effectiveness_of_argument	4.021	.046*	.018
Perceptions_on_message	2.109	.148	.010
Gender	8.237	.005**	.037
GenderRobot	.177	.675	.001
PersuasiveMessage	.573	.565	.005
Gender * GenderRobot	1.968	.162	.009
Gender * PersuasiveMessage	.907	.405	.008
GenderRobot * PersuasiveMessage	.022	.978	.000
Gender * GenderRobot * PersuasiveMessage	1.624	.199	.015

Note. $^*p < .05$, $^{**}p < .01$.
$N = 231$.