

MASTER

Investigating the Effects of Appearance Qualities on Emotional Reactions and Behavioural Tendencies towards Anthropomorphic Robots

Seyfried Cohen, Julia M.

Award date:
2023

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

**Investigating the Effects of Appearance Qualities on Emotional
Reactions and Behavioural Tendencies towards Anthropomorphic
Robots**

by J.M. Seyfried Cohen

Identity number: 1700111

In partial fulfilment of the requirements for the degree of

**Master of Science
in Human-Technology Interaction**

Supervisors:
Dr. Giulia Perugia
Dr. Stefano Guidi

Abstract

As robots become more integrated into our societies, it is important to understand whether their design affects how they are treated by people. In human-robot interactions (HRI), two separate research paths have been followed: The first, showing how humans attribute gender and age (i.e., social categories) to robots based on their appearance cues, and how these social categories elicit stereotypes, the second focusing on which stereotypes trigger emotional reactions and behavioural tendencies towards robots. The aim of this research was to connect these two paths by understanding whether the appearance qualities used to suggest gender and age in robots can themselves go as far as to elicit emotional reactions and behavioural tendencies towards them. If it would be discovered that certain appearance qualities influence people's behaviours, especially in terms of harmful behavioural tendencies, this result could inform the design of robots. Participants (N=714) were involved in an online survey, they were asked to rate the appearance qualities of one out of 70 robots from the Anthropomorphic roBOTs database (ABOT), along with their perceptions of the robot's gender, age, and stereotypical traits. Moreover, they were asked to express their emotional reactions and behavioural tendencies towards the robot. The results indicate that certain appearance qualities can have a direct effect on emotional reactions and behavioural tendencies, for instance, white hair predicting active harm. This research is important as it shows that the way we design robots can shape how people perceive and behave towards them.

Keywords: Anthropomorphic Robots, HRI, Appearance Qualities, Emotional Reactions, Behavioural Tendencies

Table of Contents

1. INTRODUCTION	5
2. RELATED WORK	8
2.1 HUMAN-HUMAN INTERACTION (HHI)	8
2.1.1 <i>Effect of appearance qualities on social categories</i>	8
2.1.2 <i>Effect of appearance qualities and social categories on stereotypical traits</i>	9
2.1.3 <i>Effect of stereotypical traits on emotional reactions and behavioural tendencies</i>	9
2.2 TRANSITIONING FROM HUMAN-HUMAN INTERACTIONS (HHI) TO HUMAN-ROBOT INTERACTIONS (HRI)	14
2.3 HUMAN-ROBOT INTERACTION (HRI).....	15
2.3.1 <i>Effect of appearance qualities on social categories</i>	15
2.3.2 <i>Effect of appearance qualities and social categories on stereotypical traits</i>	17
2.3.3 <i>Effect of stereotypical traits on emotional reactions and behavioural tendencies</i>	20
3. HYPOTHESES	22
3.1 PRE-ANALYSIS: APPEARANCE QUALITIES ON SOCIAL CATEGORIES	22
3.1.1 <i>Appearance qualities on masculinity</i>	22
3.1.2 <i>Appearance qualities on femininity</i>	23
3.1.3 <i>Appearance qualities on age</i>	23
3.2 PRE-ANALYSIS: APPEARANCE QUALITIES AND SOCIAL CATEGORIES ON STEREOTYPICAL TRAITS	24
3.3 RQ1: APPEARANCE QUALITIES ON EMOTIONAL REACTIONS.....	25
3.4 RQ2: APPEARANCE QUALITIES ON BEHAVIOURAL TENDENCIES.....	26
4. METHODOLOGY	28
4.1 STUDY DESIGN.....	28
4.2 ROBOT SELECTION.....	28
4.3 PARTICIPANTS	29
4.4 PROCEDURE AND MEASURES	30
5. RESULTS	34
5.1 DATA SCREENING.....	34
5.2 EFFECT OF APPEARANCE QUALITIES ON EMOTIONAL REACTIONS (RQ1) AND BEHAVIOURAL TENDENCIES (RQ2)	
.....	34
5.2.1 <i>Pre-analysis: Selection of appearance qualities</i>	35
5.2.2 <i>Pre-analysis: Appearance qualities on stereotypical traits</i>	39
5.2.3 <i>RQ1: Appearance qualities on emotional reactions</i>	44
5.2.4 <i>RQ2: Appearance qualities on behavioural tendencies</i>	51
6. DISCUSSION	62
6.1 INTERPRETATION OF FINDINGS	62
6.1.1 <i>Discussion pre-analysis: Appearance qualities on social categories</i>	62
6.1.2 <i>Discussion pre-analysis: Appearance qualities and social categories on stereotypical traits</i>	66
6.1.3 <i>Discussion RQ1: Appearance qualities on emotional reactions</i>	70
6.1.4 <i>Discussion RQ2: Appearance qualities on behavioural tendencies</i>	74
6.1.5 <i>General discussion</i>	78
6.2 LIMITATIONS	78
6.3 CONCLUSION	80
7. ACKNOWLEDGEMENTS	81
8. REFERENCES	82
APPENDIX	86
APPENDIX A - PARTICIPANTS NATIONALITY	86

APPENDIX B - SURVEY	89
APPENDIX C - PRE-ANALYSIS: APPEARANCE QUALITIES THAT DO NOT HAVE A SIGNIFICANT EFFECT.....	99
APPENDIX D - ROBOTS INCLUDED IN THIS STUDY	102

Investigating the Effects of Appearance Qualities on Emotional Reactions and Behavioural Tendencies towards Anthropomorphic Robots

1. Introduction

Over the years, social robots have gained a great prominence within different environments (Savela et al., 2018). Social robots can increasingly be found in care homes as companions of older people (Abdi et al., 2018), or in schools as assistants to support the education of children (You et al., 2006).

Many of the robots used in environments where human-robot interaction (HRI) occurs more frequently are anthropomorphic robots. A large group of anthropomorphic robots used in HRI are humanoid robots, robots specifically designed to be similar to humans in terms of their appearance, behaviours, and the ways in which they interact (Zhao, 2006). Humanoid robots by virtue of their design, can display a multitude of age and gender cues (Powers & Kiesler, 2006). However, what truly makes humans perceive and interact with them in a ‘human-like’ way has not only to do with the robot’s appearance or behaviour, but rather with the ability of the person to interpret and classify appearance and behaviour as such. For instance, the process of *gendering* a robot has been viewed as a two-step progression (Perugia et al., 2022). First, there is *gender encoding*, which refers to designers incorporating gender cues into the robot. Then, there is *gender decoding*, which refers to humans decoding the gender cues in the robot, and attributing gender as a result (Perugia et al., 2022). Gender decoding is seen as a spontaneous process that can occur whether a robot is purposely encoded with gender cues or not (Perugia & Lisy, 2022). For instance, people attribute gender to seemingly genderless robots such as the cleaning robot ‘Roomba’ (Marchetti-Bowick, 2009).

Similar to gender decoding happening unpredictably, gender encoding can occur subconsciously as well. Designers might not be aware of which gendered appearance qualities they are encoding into a robot, making it even more important to understand what the implications of such involuntary gender encoding is. Gender and age are both social categories, defined as categories that are “both abstract and concrete (...) concrete, in that instances are physical objects (persons) (and) abstract, in that instances are not observable, since they are fulfilling a socially defined role.” (Dahlgren, 1985). Social categorisation refers to “a cognitive process for

understanding and explaining the world, (...) (t)hat is, the groups we belong to such as genders, ethnicities, religions, and nations (...) based on social categories” (McGarty, 2018).

Aside from the effect that robot’s appearance cues have on the perception of gender and age, social categories, and the appearance qualities used to suggest them, are known to also have an effect on stereotypes (Perugia et al., 2023). Stereotypical traits generally focus on *agency* (sometimes referred to as competence) and *communion* (sometimes referred to as warmth), the former describing a robot’s perceived ability to perform tasks and achieve goals, and the latter focusing on the robot’s perceived social desirability and friendliness (Mieczkowski et al., 2019; Perugia et al., 2023). From previous research it becomes clear that appearance cues that evoke the perception of masculinity (e.g., the presence of a torso, legs, or a large shoulder width (Bernotat et al., 2021; Perugia et al., 2022)) lead to feelings of higher agency, which makes people perceive robots as more capable to perform stereotypical male tasks (e.g., being a bodyguard, shovelling the snow (Parlangeli et al., 2023)). For female robots gender stereotyping can be found too (e.g., robot ‘Nadine’ was found to be the best fit for decorating the house and staying home with a sick child (Parlangeli et al., 2023)). Understanding which stereotypical traits are evoked from appearance cues is important to understand how to design social robots in a way that reduces bias and stereotypes to the minimum.

While stereotypical traits themselves may not be inherently harmful, for instance male robots being perceived as being suitable ‘bodyguards’ (Parlangeli et al., 2023), they can influence our perception of a robot's suitability for specific roles, for instance male robots being perceived as being more capable to do most tasks (e.g., teaching, being a driver, mowing the lawn (Parlangeli et al., 2023)), compared to female robots who are perceived as being good child caretakers (Eyssel & Hegel, 2012; Parlangeli et al., 2023). Stereotypical traits thus reflect the expectations and beliefs we hold regarding the capabilities of robots in performing certain tasks.

Furthermore, stereotypical traits are known to lead to certain emotional reactions and behavioural tendencies (Cuddy et al., 2007; Fiske et al., 2002; Mieczkowski et al., 2019). Mieczkowski and colleagues (2019) looked at the effect that robot’s stereotypical traits have on human behaviours by replicating the Behaviours from Intergroup Affect and Stereotypes (BIAS) map developed by Cuddy and colleagues (2007). The researchers found that certain stereotypical traits (e.g., high communion) of robots affect how humans behave towards them (e.g., active facilitation (e.g., help, protect)). While there are helpful behaviours that can arise through stereotypical traits and emotions (e.g., admiration leading to active facilitation (e.g., help, protect)),

it is also important to consider the more harmful behavioural tendencies that can arise (e.g., envy leading to active harm (e.g., fight, attack)). While Mieczkowski and colleagues (2019) have replicated the BIAS map discovering a relationship between the stereotypical traits, emotional reactions and behavioural tendencies robots elicit, to date, the specific role that appearance cues have, in particular *qualities* of appearance cues (e.g., *long* hair, *muscular* torso, face with *make-up* etc.), on stereotypes, emotional reactions and behavioural tendencies, has not yet been studied in HRI.

Emotions such as admiration can lead to positive, helpful behaviours towards the robot (e.g., admiration leading to active facilitation (e.g., help, protect) (Cuddy et al., 2007)), whereas envy can trigger behaviours that are less positive (e.g., envy leading to active harm (e.g., fight, attack) (Cuddy et al., 2007)). Since robots will most likely become more integrated into our societies, it is important to understand whether their design affects how they are perceived and treated by people.

Through this research the aim is to: (1) identify a number of appearance qualities that could be used to describe robot's appearance beyond the mere presence of certain body parts (as in the Anthropomorphic roBOT database (ABOT); Phillips et al., 2018); (2) collect a dataset featuring people's perception of these appearance qualities, social categories, stereotypical traits, emotional reactions and behavioural tendencies for a set of 70 anthropomorphic robots from the ABOT database, and (3) identify the relationships (direct or mediated) between appearance qualities, social categories, stereotypes, emotional reactions, and behavioural tendencies. If it were discovered that certain appearance qualities influence people's behaviour, especially in negative ways (e.g., active harm), this result could inform the way we design humanoid robots, and could help pursue robot designs less prone to elicit stereotyping, bias, and discrimination in humans. Besides, it could help to understand whether the same types of biases elicited in humans on the mere basis of someone else's body also extend to humanoid robots.

2. Related Work

In this section, I will first present the research on appearance qualities, social categories, stereotypical traits, emotional reactions and behavioural tendencies in human-human interactions (HHI), before going into the latest research that has been done in human-robot interactions (HRI). Finally, I will highlight what gaps currently exist in the related work.

2.1 Human-human interaction (HHI)

2.1.1 Effect of appearance qualities on social categories

Early research in psychology has focused on understanding which human appearance qualities affect how the social categories of gender and age are perceived (George & Hole, 2000; Hehman et al., 2014). Research on the social categorisation of gender has found that facial qualities such as large eyes and rounder face-shapes are generally associated with women, whereas a strong jawline and high eyebrows are more frequently linked to men (Cunningham, 1986; Perrett et al., 1998). Nowadays products such as make-up allow people to amplify biological features (Russell, 2009). With more appearance qualities and enhanced facial features, social categorisation tends to quicken (Hehman et al., 2014). Hehman and colleagues (2014) found that appearance qualities that are aligned with each other, in other words qualities that are perceived to “go together”, allow gendering to happen faster. For instance, big eyes combined with feminine jawline (vs. big eyes combined with masculine features) lead to a faster perception of a person being perceived as female.

Not only do appearance qualities affect the perception of gender, George and Hole (2000) examined whether appearance qualities affect the perception of age as well. The researchers presented participants with different images of younger and older people’s faces, images that had either not been manipulated or had been adjusted slightly to highlight certain features. Overall, researchers found that participants were able to accurately determine age, and that manipulations did not greatly affect the accuracy of age perception. George and Hole (2000) discovered that the appearance qualities of face shape and “surface-patterns”, which include wrinkles and skin pigmentations, give us the most information about a person’s age.

2.1.2 Effect of appearance qualities and social categories on stereotypical traits

Besides the effects that appearance qualities have on social categories of gender and age, research has also investigated the role that social categories have on eliciting stereotypes (Hummert et al., 1997; Kaufmann et al., 2016). Hummert and colleagues (1997) asked participants to evaluate images of older adults and found that in the group of older women, those that were the youngest were stereotyped more negatively (e.g., bitter, stubborn) than their male counterparts. Negative stereotypes only increased for men that were the oldest in the sample. Hence, Hummert and colleagues (1997) found a mixed effect of social categories on stereotyping (i.e., younger older women being stereotyped more negatively than younger older men). They found that this mixed effect was mitigated through facial expressions, that is, friendly facial expressions of older women received more positive stereotypes (e.g., kind, loving). Another example where social categories have been seen to induce stereotypes was shown by Kaufmann and colleagues (2016). The researchers asked participants to decide whether or not they would hire a fictitious job applicant. In one condition, participants could read the age in the participant's profile, in another condition participants viewed a picture of the fictitious candidate without knowing their precise age. The researchers discovered that stereotypes of bad health as well as lower stamina were triggered in the condition where participants viewed a picture of the job applicant. Based on the negative impressions that were formed, older fictitious applicants had a lower chance of being hired.

2.1.3 Effect of stereotypical traits on emotional reactions and behavioural tendencies

Research has not only investigated what effects appearance qualities and social categories have on stereotypes, but also the implications of stereotypes on emotional reactions and behavioural tendencies. In the following, I will introduce the Stereotype Content Model (SCM) and the Behaviours from Intergroup Affect and Stereotypes (BIAS) map. The focus here is not to replicate these two, this has already been done in HRI, but rather to understand the relationships between stereotypical traits, emotional reactions, and behavioural tendencies, and by doing so, I can then expand upon this knowledge and establish the relationship between appearance qualities on emotional reactions and behavioural tendencies, considering the influence of social categories and stereotypical traits in the process.

2.1.3.1 The Stereotype Content Model

Fiske and colleagues (2002) developed the ‘Stereotype Content Model’ (SCM), which showed how different social groups, for example rich people, housewives or old people elicit specific stereotypes and how these stereotypes lead to specific emotional reactions. A social group "constitutes a set of two or more persons who are linked through interaction", that share certain characteristics and a common identity (Biddle, 1979). The shared characteristics of a social group allows others (i.e., ‘out-group’ members of other social groups or even ‘in-group’ members of the same social group) to form stereotypes of the social group.

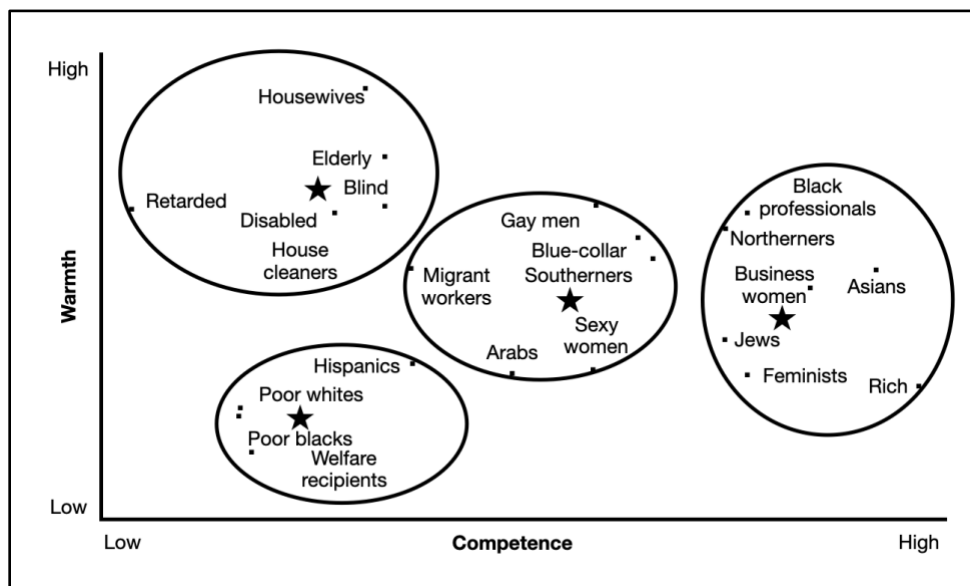
Social perception, especially from an evolutionary perspective, is crucial to determine whether someone can be deemed a “friend or foe” (Fiske et al., 2002; Reeves et al., 2020). The researchers argue that when we meet someone for the first time, we first focus on their intentions, their *warmth* and later decide whether they have the *competence* to act upon these intentions. Across different cultures and times, the dimensions of warmth (sometimes referred to as friendliness or helpfulness) and competence (sometimes referred to as skilfulness or creativity) have appeared consistently (Reeves et al., 2020). Important to note is that the dimensions of competence and warmth are not viewed as conscious evaluations that people make of one another, but rather are automatic and unconscious judgements (Eyssel & Hegel, 2012).

The SCM proposes that unique combinations of warmth and competence can either lead to *univalent* stereotypes, where both warmth and competence are low or high or *ambivalent* (or mixed) stereotypes, where warmth overpowers competence or vice versa (Fiske et al., 2002). In total, there are four different possible combinations of the warmth and competence dimensions; the univalent stereotypes (low warmth/ low competence; high warmth/ high competence) and the ambivalent stereotypes (high warmth/ low competence; low warmth/ high competence). Based on these combinations, different emotions towards social groups are elicited. Fiske and colleagues (2002) argue that there are four emotional reactions that can arise: pity, envy, admiration and contempt. For instance, rich people are perceived as being low in warmth but high in competence, and therefore can elicit feelings of envy (Figure 1). Vice versa, old people are perceived as high in warmth but low in competence, which leads to feelings of pity (Figure 1). Fiske and colleagues (2002) pointed out the previous research has generally focused on negative univalent stereotypes, whereas the SCM also looks into positive univalent stereotypes and more importantly, ambivalent stereotypes that are more functionally consistent with the views that society has of social groups, such as old people being perceived as high in warmth but low in competence.

The researchers predicted that all social groups would fit within these different combinations of warmth and competence, which could be graphically represented by four quadrants that have warmth and competence, going from low to high, as the axes in the SCM (Figure 1).

Figure 1

The SCM - Four cluster solutions to combinations of the dimensions of warmth and competence (Fiske et al., 2002).



2.1.3.2 The Behaviours from Intergroup Affect and Stereotypes (BIAS) map

The Behaviours from Intergroup Affect and Stereotypes (BIAS) map builds upon the SCM (Cuddy et al., 2007). The BIAS map is a framework that shows how stereotypical traits not only elicit emotional reactions but rather can be seen as triggering behavioural tendencies towards social groups as well. The researchers hypothesised that behaviours towards social groups (referred to as target group) can be either *active*, direct efforts to affect the target group, or *passive*, less direct efforts to affect the target group, *facilitation*, favourable outcomes for the target group, or *harm*, detrimental outcomes for the target group. All these qualities of behaviours lead to the combination of four main possible behavioural tendencies in the BIAS map: *active facilitation*, which means to act *for* a target group with the intention to benefit the group (e.g., helping, defending), *active harm*, which means to act *against* a target group with the intention to harm the group (e.g., harassment,

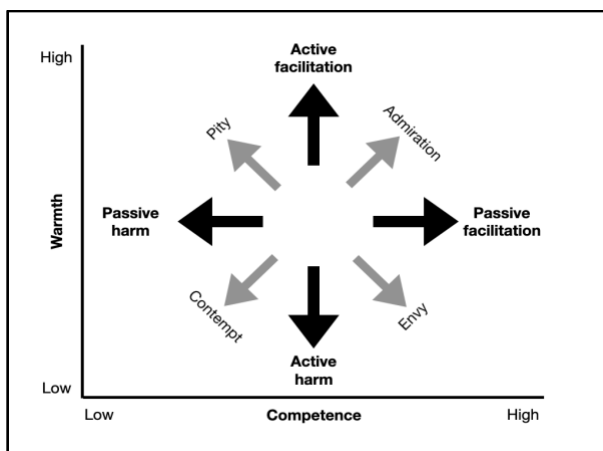
hate crime), *passive facilitation*, which means to act *with* a target group for one's own purpose (e.g., working with a member of the target group), and *passive harm*, which means to act *without* a target group (e.g., avoiding eye contact with the target group or limiting the target group's access to resources, see Figure 2).

Furthermore, Cuddy and colleagues (2007) used the BIAS map to show which stereotypes could determine whether a behavioural tendency would be active, passive, facilitative or harmful. They argued that the warmth dimension would predict active behaviours as well as the valence of the behavioural tendency (i.e., active facilitation or active harm), whereas the competence dimension would determine passive behaviours and the valence of these passive behavioural tendencies (i.e., passive facilitation or passive harm) (Figure 2).

Similar to the SCM, the BIAS map discloses that there are four emotional reactions that originate from the dimensions of warmth and competence: pity, admiration, envy and contempt. New contributions from the BIAS map are, however, that emotional reactions can also elicit behavioural tendencies. While we know from the SCM that each combination of competence and warmth leads to a unique emotional reaction, for instance, high-warmth and low-competence eliciting feelings of pity (Fiske et al., 2002), Cuddy and colleagues (2007) argued that each of these emotional reactions also trigger two behavioural tendencies. Pity can lead to active facilitation or passive harm, envy can lead to active harm or passive facilitation, admiration can lead to active or passive facilitation, and contempt can lead to active and passive harm (Figure 2).

Figure 2

The BIAS map is composed of two-dimensional stereotypes (competence and warmth), four emotional reactions (pity, admiration, envy, and contempt) and four behavioural tendencies (active facilitation, active harm, passive facilitation, and passive harm) (Cuddy et al., 2007).



Cuddy and colleagues (2007) tested this framework by running four studies. The first study was a telephone survey, where participants were asked about their perception of common social groups within society. They selected five groups to represent each of the four quadrants of the competence-warmth space (i.e., five groups to represent the quadrant of low warmth-low competence), leading to 20 social groups that participants were asked to evaluate. Cuddy and colleagues (2007) asked participants about stereotypical traits, emotional reactions as well as their behavioural tendencies towards social groups to understand the relationship that exists between stereotypical traits and behaviours as well as emotions and behaviours. To reduce the social desirability as much as possible, they asked participants to rate how these groups were “perceived by Americans”. For instance, for behavioural tendencies, Cuddy and colleagues (2007) asked the same question to participants, each time replacing the social group and behavioural tendency for different questions:

“(...) I am going to ask you about the ways people in America generally behave toward [group] as a group? Do people tend to [behaviour, e.g., help] [group]?”

In the second and third study, researchers looked at the causal links between stereotypes and behavioural tendencies as well as between emotions and behavioural tendencies. Study two and three were similar to the first one, with the only difference being that either competence or warmth was chosen as a stereotype and, in the third study, admiration, contempt, envy or pity was used to describe the social group. All questions were asked in a similar way to the example question shown above. In their final study, Cuddy and colleagues (2007) examined the relationship between anger and fear in the BIAS map. However, to stay consistent with emotions studied in the SCM and emotions studied in human-robot interactions (HRI), the emotions that will be examined in this thesis are pity, envy, admiration, and contempt.

Cuddy and colleagues (2007) found that stereotypical traits not only elicit emotional reactions, but rather impact behavioural tendencies, in-line with the BIAS map. The warmth dimension was found to elicit active behavioural tendencies. For instance, more perceived warmth of a social group led to an increase in active facilitation (e.g., helping) and reduced active harm (e.g., harassing). The competence dimension predicted passive behavioural tendencies, with more perceived competence of a social group increasing passive facilitation (e.g., associating) and reducing passive harm (e.g., neglecting).

Aside from the dimensions of warmth and competence affecting behaviours, Cuddy and colleagues (2007) found that each of the emotions triggered behavioural tendencies as well. Social groups that were admired (i.e., high in warmth and high in competence) elicited active facilitation and passive facilitation. However, hated social groups (i.e., low in warmth and low in competence) elicited active and passive harmful behavioural tendencies. Groups that were envied (i.e., high in competence and low in warmth) led to passive facilitation and active harm, whereas pitied groups (i.e., low in competence and high in warmth) led to active facilitation and passive harm. An important concept they describe in their research is “emotional priority”, which means that emotions have a greater effect on behavioural tendencies than stereotypes do.

2.2 Transitioning from human-human interactions (HHI) to human-robot interactions (HRI)

Previous research has found that humans generally respond and interact with robots in similar social ways and with similar social processes as they do when interacting with other people (Reeves et al., 2020). Across different studies, research has found proof that social characteristics in HHI are replicated in HRI (Reeves et al., 2020): people are as friendly and polite towards computers as they are towards people (Nass, 2004), introverted and extroverted ‘text-to-speech’ voices programmed into computers lead to impressions similar as in HHI (Nass & Lee, 2001), and there are different levels of physiological arousal of people when asked to touch different mechanical parts of robots (Li et al., 2017). Another important similarity between HHI and HRI is the speed with which people form judgements about robots, which often occurs within 33-38 milliseconds after viewing the face of the robot (Bar et al., 2006).

What is known thus far from previous research in HHI is that appearance qualities can influence social categorisation, and that belonging to a social group shapes the way we perceive that group in terms of stereotypes, emotional reactions and behavioural tendencies. The next section will dive into what is known about the relationship between appearance qualities, social categories, stereotypical traits, emotional reactions and behavioural tendencies in HRI, and which research gaps currently exist in HRI.

2.3 Human-robot interaction (HRI)

2.3.1 Effect of appearance qualities on social categories

First, I start by looking at research that has studied the effects of appearance qualities on social categories in HRI. This thesis does not focus on the overarching effects of appearance cues (e.g., presence of a torso, hair, face), but rather dives deeper into the specific *qualities* of these appearance cues that trigger social categorization (e.g., *muscular* torso, *long* hair, *face with make-up* etc.). Here I report the findings from previous research in HRI that discuss which qualities of appearance cues elicit social categorisation.

2.3.1.1 Appearance qualities on masculinity and femininity

Eyssel and Hegel (2012) studied the role that facial gender cues have on the social categorisation of robots as male or female. In their research, they presented participants with two sets of humanoid robot images that were similar to each other, the only difference between them being its hair style and the shape and colour of the lips. The first robot had longer hair and fuller red lips, whereas the second robot had shorter hair and thinner lips. The researchers found that while long-hair elicited female categorisation, short-hair triggered male categorisation.

Similarly, Bernotat and colleagues (2021) were interested in understanding the effects that specific appearance qualities have on how people perceive the gender of robots. They studied the effects different waist-to-hip ratios, as well as shoulder widths, would have on a robot's perceived gender. Participants in the study were presented with images of two humanoid robots, with different (i.e., larger, or smaller) waist-to-hip and shoulder width ratios and were asked in a questionnaire to rate whether the body shape of the robots was more feminine or masculine. The researchers found that a large waist-to-hip ratio lead participants to perceive the robot as more masculine. They also discovered that shoulder width affected the perception of robots' gender, with a wider shoulder width leading to a more masculine gender attribution.

Furthermore, Perugia and colleagues (2022) examined which appearance cues affected the attribution of age and gender to a robot. In their survey, they divided the 251 images of robots from the Anthropomorphic roBOTS (ABOT) dataset (Philips et al., 2018) into five groups and randomly allocated participants to these groups. Participants were then asked to attribute age and gender to the robots. The appearance cues they investigated were derived from the ABOT database scores (Philips et al., 2018): *surface look* features, which rate the presence in a robot of eyelashes, head hair, skin, genderedness, nose, eyebrows, and apparel, *body manipulators*, which rate the

presence in a robot of arms, a torso, fingers, and legs, and *facial features*, which rate the presence in a robot of a face, eyes, a head and a mouth. The researchers found that body manipulators were the most important predictors of masculinity.

In a similar way to the appearance qualities that affect the perceived masculinity of a robot, there are also appearance qualities that influence the perceived femininity of a robot. Kalegina and colleagues (2018) examined which facial qualities affect the perceived age and gender of a robot. In their research, they presented participants with a set of 12 images of robots that ranged anywhere from detailed humanoid robots with detailed facial features, to non-humanoid robots with few facial features, such as different ‘skin’ colours, eyebrows, noses, and eye shapes. Participants were then asked to rate the images of the robots, amongst other criteria, on their perceived gender and age. The researchers found that a strong predictor of a robot being perceived as female was the colour of the face. White faces were ranked far more feminine than masculine, which they argued was the result of women being biologically more prone to have lighter skin than men.

Furthermore, Perugia and colleagues (2022) discovered that the strongest predictors of femininity were surface look features. The researchers found that the presence of apparel, long eyelashes, long hair and full and red lips are known to lead to the perception of higher femininity (Eyssel & Hegel, 2012; Perugia et al., 2022). While short hair and thin lips are associated with male robots, long hair and full red lips are linked to female robots (Eyssel & Hegel, 2012). Here, an important finding was that robots with red lips (i.e., the robot wearing make-up) were perceived as being far more feminine than robots without red lips (Eyssel & Hegel, 2012).

2.3.1.2 Appearance qualities on age

Appearance qualities have also been shown to influence the perception of robots’ age. Kalegina and colleagues (2018) found in their set of 12 robots that the ones with a smaller eye distance (i.e., big eyes), a nose and hair were perceived as more childlike than the others. Hence, the presence of these features negatively predicts the age of robots. These findings are in line with observations made in the field of biology, with the so-called “baby schema” (Lorenz, 1943). In 1943, the zoologist Konrad Lorenz explained that humans tend to find common features of infants appealing and “cute”, regardless of whether these infants are humans or animals - referred to as the “baby schema”. Some of these common features include a large head, big eyes and overall rounder bodies. Perugia and colleagues (2022) found that facial features of a robot generally lead

to the impression that the robot is younger in age, which might relate to the baby schema as most of the defining elements of the baby schema are located on the face (e.g., big eyes, bulgier head).

2.3.2 Effect of appearance qualities and social categories on stereotypical traits

The effect of appearance qualities and social categories on stereotypical traits have been studied in HHI (Cuddy et al., 2007; Fiske et al., 2002), but also in HRI (Bernotat et al., 2021; Calvo-Barajas et al., 2020; Eyssel et al., 2012; Perugia et al., 2023). While in the previous section, I presented the related work on the appearance qualities that have an effect on the perception of masculinity, femininity, and age in a robot, here I will introduce the literature on the effects of appearance qualities and social categories on stereotypical traits, namely communion and agency. I want to highlight here that while some research uses the related stereotypical traits of *warmth* and *competence*, there are slight differences with these to *communion* and *agency*. While warmth and communion are very similar concepts, competence refers more to *potential* action whereas agency refers to *actual* action and capabilities (Cuddy et al., 2008), which is why I prefer to use *agency* and *communion* in this thesis. In the following, I will report the stereotypical traits that the research used (e.g., warmth and competence), however, in this thesis the stereotypical traits of communion and agency will be examined.

2.3.2.1 Appearance qualities (that predict femininity) and femininity on communion

Research has investigated which social categories elicit the stereotypical traits communion and agency in humanoid robots. In an within-subjects design, Eyssel and Hegel (2012) wanted to understand which stereotypical traits would be attributed to male and female robots. They gave participants a list of 12 adjectives for communion (e.g., affable, friendly, polite, affectionate) and another 12 for agency (e.g., assertive, dominant, determined, authoritative), and asked them to evaluate the robots based on these. The researchers found that female robots received higher scores for communion than male robots did.

Furthermore, Bernotat and colleagues (2021) looked at the effects of appearance qualities on the perception of the robot's gender, and examined whether the perceived gender evoked certain stereotypical traits. They asked participants to rate the robots based on 14 adjectives, seven for agency (e.g., assertive, authoritative, confident, determined) and seven for communion (e.g., friendly, empathetic, sensitive, affectionate), similar to Eyssel and Hegel (2012). They found that the male robot was perceived as being agentic, in-line with their predictions.

Additionally, using the Stereotype Content Model (SCM) developed by Fiske and colleagues (2002), that describes the content and structure of stereotypes, Perugia and colleagues (2023) examined the effect that appearance qualities, gender and age have on stereotypes of communion and agency. Eighty images of robots from the ABOT dataset (Philips et al., 2018) were randomly divided into four groups and participants were randomly allocated to one of these groups. Participants were then asked to evaluate the robot in terms of agency and communion, as well as based on how suitable for female and male tasks they perceived the robot to be. To estimate the appearance of the robot, the researchers used the scores from the ABOT dataset (Philips et al., 2018), which measure the presence or absence of in a robot of body manipulators, surface look, and facial features. In their research, multiple effects of appearance qualities and social categories on communion were found. The most significant finding was that surface look features (e.g., long eyelashes, long hair) had a positive effect on communion mediated by the robot's perceived femininity (Perugia et al., 2023). Unfortunately, these effects were only marginally significant, probably due to the inferior number of feminine robots in the ABOT database.

Furthermore, research has studied what effects gender and facial qualities of a robot (e.g., upwards eyebrows and lips) have on trust and psychological reactance in humans (Ghazali et al., 2018). In a study, Ghazali and colleagues (2018) asked participants to play an online trust game, and used the 'SociBot™' robot, which was able to display different facial qualities, to persuade participants to make certain choices while playing the game. The researchers changed the facial qualities of the robots and found that the robot was trusted and its advice was followed more when it had upward eyebrows and lips. The gender of the robot did not influence the trust that participants felt towards it. One of the most interesting findings, however, was that liking the robot fully mediated any direct effects of facial qualities on trust and reactance.

Liking also played a big role in a study by Calvo-Barajas and colleagues (2020), who found that female robots, when embodying the stereotypical trait communion in the form of emotions, are liked more. Calvo-Barajas and colleagues (2020) looked at the role that facial expressions displayed by gendered robots have on the emotional reactions of children. They wanted to examine which facial expressions influenced the trustworthiness and likability of the robot by testing how a range of anger and happiness emotions with different valences and intensities were perceived by children. The young participants were presented with two robots, one male-like and one female-like robot and were asked to rate the facial expressions they saw. The researchers discovered that there was an interaction effect between the perceived gender of the robot and intensity of the

emotions it displayed on the robot's likability. While male robots were liked more when they displayed high anger, female robots were liked more when displaying positive and subtle emotions (e.g., low intensity happiness). These results show how female robots are liked better when they abide by communion stereotypes and act friendly (Guidi et al., 2022; Perugia et al., 2023). In this thesis too I expect to find that appearance qualities that predict femininity along with femininity evoke the stereotypical trait communion.

2.3.2.2 Appearance qualities (that predict masculinity) and masculinity on agency

Appearance qualities and social categories have also been found to have an effect on stereotypical traits of agency. While Eyssel and Hegel (2012) found that female robots received higher scores of communion, they found that male robots were perceived as having greater agency. Similarly, Bernotat and colleagues (2021) found in their research too that the stereotypical trait agency was attributed to male robots. I therefore argue that masculinity evokes the stereotype of agency. I too predict to come to similar conclusions, that appearance qualities that predict masculinity and masculinity itself will lead to the stereotypical trait agency.

2.3.2.3 Appearance qualities (that predict age) and age on communion

The effect of appearance qualities (that predict age) on communion was found by Perugia and colleagues (2023). The researchers discovered that facial features (e.g., big eyes, nose) have a positive effect on communion, which was mediated by the robot's perceived age. Dunstan and Hoffman (2023) argue that robots are purposely designed to look "cute". They argue that cuteness conveys a bit of *naïveté* and helplessness, but at the same time also vulnerability and loyalty. A robot being perceived as child-like and "cute" shapes the expectations that people have of the robot, as a child-like robot may be perceived as being less capable of doing certain tasks compared to the perceived capabilities of an adult-like robot, and instead elicit feelings of communion and warm, caring feelings (Dunstan & Hoffman, 2023). Interacting with a robot that is more child-like makes the user interact with it in a more "tender and caring way" and allows for the user to become somewhat of a caretaker and to "treat it like an infant" (Breazeal & Foerst, 1999). Here, I too expect to find that appearance qualities that predict age along with age to elicit the stereotypical trait communion.

2.3.3 Effect of stereotypical traits on emotional reactions and behavioural tendencies

2.3.3.1 The Stereotype Content Model (SCM)

Up until now the effects that appearance qualities and social categories have on stereotypical traits have been discussed. The SCM, however, shows which effects the dimensions of warmth and competence have on emotional reactions.

There is only limited research that has specifically tried to replicate the SCM in HRI. For instance, Perugia and colleagues (2023) followed the same structure as the SCM and used the SCM to cluster their robots based on the valence of the dimensions of communion and agency. Similar to what found by Fiske and colleagues (2002) and others after them, a four-cluster solution was found to be the most ideal also for social robots in both papers: cluster 1 (low agency/ low communion), cluster 2 (low agency/ high communion), cluster 3 (high agency/ low communion), and cluster 4 (high agency/ high communion). From Mieczkowski and colleagues (2019), who by replicating the BIAS map also replicated the findings of the SCM, found that groups high in communion and agency elicit feelings of admiration, low agency/ high communion elicits feelings of pity, however not that high agency/ low communion elicits feelings of envy nor that low agency/ low communion elicits feelings of contempt.

2.3.3.2 The Behaviours from Intergroup Affect and Stereotypes (BIAS) map

In HRI, Mieczkowski and colleagues (2019) were the first to study how stereotypes are not only connected to people's emotional reactions but rather behavioural tendencies towards robots. In an online survey, they asked participants to evaluate a set of 342 social robots from the Stanford Social Robot Database. They asked 10 participants to look at one of the 342 robot images and answer a set of questions regarding the robot's perceived warmth and competence (i.e., stereotypical traits), the envy, admiration, pity and contempt it elicited (i.e., emotional reactions), and the active facilitation, active harm, passive facilitation, and passive harm participants felt towards it.

The structure of each question was similar to the questions posed by Cuddy and colleagues (2007). The researchers asked participants, for instance:

“How [warm, tolerant, good-natured, sincere] is this robot?” and “How [competent, confident, independent, competitive, intelligent] is this robot?”

Mieczkowski and colleagues (2019) found partial proof that stereotypical traits warmth and competence affect emotional reactions (as was shown by Cuddy and colleagues (2007)). As mentioned above, while high competence/ high warmth groups lead to more admiration and for low competence/ high warmth groups to more pity, researchers did not find that high competence/ low warmth lead to envy, nor that low competence/ low warmth lead to more contempt. Furthermore, the researchers found partial proof of their hypothesis that stereotypes impact behavioural tendencies. The research showed that high warmth predicts active facilitation and high competence predicts passive facilitation. However, there seemed to be a weaker relationship between clusters of social robots and predicted emotions than indicated in previous work of the BIAS map. The researchers did not find proof that low warmth predicts active harm, nor that low competence predicts passive harm.

Mieczkowski and colleagues (2019) found, in-line with the BIAS map, that each of the four emotional reactions predicts two behavioural tendencies. Admiration was found to predict active and passive facilitation, contempt was found to predict active and passive harm, envy was found to predict active harm and passive facilitation, and pity was found to predict active facilitation and passive harm. They also found the concept of “emotional priority” in their research, as emotions were better predictors of behavioural tendencies than stereotypes were.

From this research it becomes clear how the BIAS map can be applied in HRI, as to say that it’s known how the stereotypes robots elicit in terms of warmth and competence affect people’s emotional reactions and behavioural tendencies. However, what is not yet known is whether the appearance qualities of robots, and the social categories these elicit, have a role in prompting stereotypical trait attributions, as well as emotional reactions and behavioural tendencies. This is precisely what I attempt to do in this thesis. As such I will focus on the following research questions:

- *RQ1: To what extent do **appearance qualities**, and the social categories and stereotypical traits they elicit, affect people’s **emotional reactions** towards anthropomorphic robots?*
- *RQ2: To what extent do **appearance qualities**, and the social categories and stereotypical traits they elicit, affect people’s **behavioural tendencies** towards anthropomorphic robots?*

3. Hypotheses

In the following section, I present the hypotheses that will be tested in this thesis, that derive from the previous literature. The hypotheses will be listed starting with (1) the effects of appearance qualities on social categories (gender and age), continuing with (2) the effects of appearance qualities and social categories on stereotypical traits (communion and agency), and finishing with (3) the effects of appearance qualities on emotional reactions, and (4) appearance qualities on behavioural tendencies. The overall aim of this thesis is twofold: (i) determine the link between robots' appearance qualities and people's emotional reactions and behavioural tendencies towards them, and (ii) understand whether social cues (perceived gender and age) and stereotypical traits (communion and agency) have a mediating role in this relationship.

While previous research in HHI often looked at the stereotypical traits warmth and competence, this thesis will focus on communion and agency, which is consistent with previous research in HRI (Abele & Wojciszke, 2007; Bernotat et al., 2021, Eyssel & Hegel, 2012; Perugia et al., 2022; Perugia et al., 2023)

3.1 Pre-analysis: Appearance qualities on social categories

3.1.1 Appearance qualities on masculinity

Since previous literature in HRI showed that short hair (Eyssel & Hegel, 2012), a large waist-to-hips ratio and wide shoulder width (Bernotat et al. 2021), and the presence of body manipulators (i.e., arms, torso, fingers, hands, and legs (Philips et al., 2018)) (Perugia et al., 2022) affect the perception of masculinity, I hypothesise that:

H1a: The presence of short hair positively predicts masculinity.

H1b: The presence of a torso (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, curvy, or with visible mechanics) predicts masculinity.

H1c: The presence of arms (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics) predicts masculinity.

H1d: The presence of hands (e.g., moveable, or able to grasp objects) predicts masculinity.

H1e: The presence of fingers (e.g., long, short, or chubby) predicts masculinity.

H1f: The presence of legs (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics) predicts masculinity.

3.1.2 Appearance qualities on femininity

Furthermore, previous literature in HRI found that white face (Kalegina et al., 2018), red, full lips (Eyssel & Hegel, 2012), apparel, long eyelashes, long hair and the presence of surface look features (i.e., eyelashes, head hair, skin, nose, eyebrows, and apparel (Philips et al., 2018)) (Perugia et al., 2022) affect the perception of femininity, hence, I hypothesise that:

H1g: The presence of skin (e.g., metallic, realistic (i.e., life-like, human-like), black, brown, white, pink, yellow, or another colour) predicts femininity.

H1h: The presence of a face with make-up (e.g., red lips) positively predicts femininity.

H1i: The presence of full lips positively predicts femininity.

H1j: The presence of apparel (e.g., t-shirt, skirt, trousers, dress, jacket, suit, mask, helmet, armour, tight, loose, formal, informal, pink, blue, dark, white, or another colour) predicts femininity.

H1k: The presence of eyelashes (e.g., long, short) predicts femininity.

H1l: The presence of head hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like) blond/ yellow, dark, grey, white, or another colour) predicts femininity.

H1m: The presence of eyebrows (e.g., thick, thin, metallic, realistic (i.e., life-like, human-like), straight, curved, raised, or frowning) predicts femininity.

3.1.3 Appearance qualities on age

Moreover, previous literature in HRI found that big eyes, the presence of a nose, the presence of hair (Kalegina et al., 2018), and the presence of facial features (i.e., face, eyes, head, and mouth (Philips et al., 2018)) (Perugia et al., 2022) affect the perception of age, hence, I hypothesise that:

H1n: The presence of eyes (e.g., big, small, rounded, elongated, expressive, moveable, metallic, realistic (i.e., life-like, human-like), mechanical, or with big pupils) predicts age.

H1o: The presence of a nose (e.g., long, short, rounded, pointed, narrow, wide) predicts age.

H1p: The presence of hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like), blond/ yellow, dark, grey, white, or another colour) predicts age.

H1q: The presence of a head (e.g., big, small, elongated, short, rounded, squared, spherical, oval, realistic (i.e., life-like, human-like), mechanical, a screen/ tablet, or a camera) predicts age.

H1r: The presence of a face (e.g., realistic (i.e., life-like, human-like), metallic, mechanical, chubby, sunken, expressive, wrinkled, bearded / with moustache, or with make-up) predicts age.

H1s: The presence of a mouth (e.g., big, small, squared, smiling, minimalistic, moveable, red, featuring full lips, or featuring thin lips) predicts age.

3.2 Pre-analysis: Appearance qualities and social categories on stereotypical traits

From previous literature in HRI, it is known that masculinity predicts agency (Bernotat et al., 2021; Eyssel & Hegel, 2012; Perugia et al., 2023), femininity predicts communion (Bernotat et al., 2021; Eyssel & Hegel, 2012; Perugia et al., 2023), and age predicts communion (Dunstan & Hoffman, 2023; Perugia et al., 2023). Here, I expect to find that the appearance qualities that elicit social categories, will in turn affect the associated stereotypical traits (communion and agency). The hypotheses are as follows:

H2a: Appearance qualities that predict masculinity (i.e., the presence of short hair, the presence of a torso (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, curvy, or with visible mechanics), the presence of arms (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics), the presence of hands (e.g., moveable, or able to grasp objects), the presence of fingers (e.g., long, short, or chubby), and the presence of legs (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics)) have a positive direct effect on agency and a positive indirect effect on agency mediated by masculinity.

H2b: Appearance qualities that predict femininity (i.e., the presence of skin (e.g., metallic, realistic (i.e., life-like, human-like), black, brown, white, pink, yellow, or another colour), the presence of a face with make-up (e.g., red lips), the presence of full lips, the presence of apparel (e.g., t-shirt, skirt, trousers, dress, jacket, suit, mask, helmet, armour, tight, loose, formal, informal, pink, blue, dark, white, or another colour), the presence of eyelashes (e.g., long, short), the presence of head hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like) blond/ yellow, dark, grey, white, or another colour), and the presence of eyebrows (e.g., thick, thin, metallic, realistic (i.e., life-like, human-like), straight, curved, raised, or frowning)) have a positive direct effect on communion and a positive indirect effect on communion mediated by femininity.

H2c: Appearance qualities that predict age (i.e., the presence of eyes (e.g., big, small, rounded, elongated, expressive, moveable, metallic, realistic (i.e., life-like, human-like), mechanical, or

with big pupils), the presence of a nose (e.g., long, short, rounded, pointed, narrow, wide), the presence of hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like), blond/ yellow, dark, grey, white, or another colour), the presence of a head (e.g., big, small, elongated, short, rounded, squared, spherical, oval, realistic (i.e., life-like, human-like), mechanical, a screen/ tablet, or a camera), the presence of a face (e.g., realistic (i.e., life-like, human-like), metallic, mechanical, chubby, sunken, expressive, wrinkled, bearded / with moustache, or with make-up), and the presence of a mouth (e.g., big, small, squared, smiling, minimalistic, moveable, red, featuring full lips, or featuring thin lips)) have a positive direct effect on communion and a positive indirect effect on communion mediated by age.

3.3 RQ1: Appearance qualities on emotional reactions

Furthermore, from previous literature in HHI (Cuddy et al., 2007) and HRI (Mieczkowski et al., 2019) it is known that groups high in competence and warmth elicit feelings of admiration, low competence/ high warmth elicits feelings of pity, high competence/ low warmth elicits feelings of envy, and low competence/ low warmth elicits feelings of contempt. Here, I anticipate that the appearance qualities effect social categories, which elicits the associated stereotypical trait (i.e., communion or agency). This will, in turn, then go on to trigger the respective emotional responses tied to the stereotype. Hence, the hypotheses are the following:

H3a: Appearance qualities that predict masculinity (i.e., the presence of short hair, the presence of a torso (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, curvy, or with visible mechanics), the presence of arms (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics), the presence of hands (e.g., moveable, or able to grasp objects), the presence of fingers (e.g., long, short, or chubby), and the presence of legs (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics)) have a positive direct effect on envy and admiration and a positive indirect effect on envy and admiration mediated by agency and masculinity.

H3b: Appearance qualities that predict femininity (i.e., the presence of skin (e.g., metallic, realistic (i.e., life-like, human-like), black, brown, white, pink, yellow, or another colour), the presence of a face with make-up (e.g., red lips), the presence of full lips, the presence of apparel (e.g., t-shirt, skirt, trousers, dress, jacket, suit, mask, helmet, armour, tight, loose, formal, informal, pink, blue,

dark, white, or another colour), the presence of eyelashes (e.g., long, short), the presence of head hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like) blond/ yellow, dark, grey, white, or another colour), and the presence of eyebrows (e.g., thick, thin, metallic, realistic (i.e., life-like, human-like), straight, curved, raised, or frowning)) have a positive direct effect on pity and admiration and a positive indirect effect on pity and admiration mediated by communion and femininity.

H3c: Appearance qualities that predict age (i.e., the presence of eyes (e.g., big, small, rounded, elongated, expressive, moveable, metallic, realistic (i.e., life-like, human-like), mechanical, or with big pupils), the presence of a nose (e.g., long, short, rounded, pointed, narrow, wide), the presence of hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like), blond/ yellow, dark, grey, white, or another colour), the presence of a head (e.g., big, small, elongated, short, rounded, squared, spherical, oval, realistic (i.e., life-like, human-like), mechanical, a screen/ tablet, or a camera), the presence of a face (e.g., realistic (i.e., life-like, human-like), metallic, mechanical, chubby, sunken, expressive, wrinkled, bearded / with moustache, or with make-up), and the presence of a mouth (e.g., big, small, squared, smiling, minimalistic, moveable, red, featuring full lips, or featuring thin lips)) have a positive direct effect on pity and admiration and a positive indirect effect on pity and admiration mediated by communion and age.

3.4 RQ2: Appearance qualities on behavioural tendencies

Building on this, I anticipate that once the appearance qualities (through social categories and stereotypical traits) elicit emotional reactions, the emotional reactions will lead to behavioural tendencies. From previous literature in HHI (Cuddy et al., 2007) and HRI (Mieczkowski et al., 2019) it is known that admiration leads to active and passive facilitation, contempt leads to active and passive harm, envy predicts active harm and passive facilitation, and pity leads to active facilitation and passive harm. I hypothesise that:

H4a: Appearance qualities that predict masculinity (i.e., the presence of short hair, the presence of a torso (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, curvy, or with visible mechanics), the presence of arms (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics), the presence of hands (e.g., moveable, or able to grasp objects), the presence of fingers (e.g., long, short, or

chubby), and the presence of legs (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics)) have a positive direct effect on active and passive facilitation, and active harm and a positive indirect effect on active and passive facilitation, and active harm mediated by admiration, envy, agency, and masculinity.

H4b: Appearance qualities that predict femininity (i.e., the presence of skin (e.g., metallic, realistic (i.e., life-like, human-like), black, brown, white, pink, yellow, or another colour), the presence of a face with make-up (e.g., red lips), the presence of full lips, the presence of apparel (e.g., t-shirt, skirt, trousers, dress, jacket, suit, mask, helmet, armour, tight, loose, formal, informal, pink, blue, dark, white, or another colour), the presence of eyelashes (e.g., long, short), the presence of head hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like) blond/ yellow, dark, grey, white, or another colour), and the presence of eyebrows (e.g., thick, thin, metallic, realistic (i.e., life-like, human-like), straight, curved, raised, or frowning)) have a positive direct effect on active facilitation, passive harm and passive facilitation and a positive indirect effect on active facilitation, passive harm and passive facilitation mediated by pity, admiration, communion, and femininity.

H4c: Appearance qualities that predict age (i.e., the presence of eyes (e.g., big, small, rounded, elongated, expressive, moveable, metallic, realistic (i.e., life-like, human-like), mechanical, or with big pupils), the presence of a nose (e.g., long, short, rounded, pointed, narrow, wide), the presence of hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like), blond/ yellow, dark, grey, white, or another colour), the presence of a head (e.g., big, small, elongated, short, rounded, squared, spherical, oval, realistic (i.e., life-like, human-like), mechanical, a screen/ tablet, or a camera), the presence of a face (e.g., realistic (i.e., life-like, human-like), metallic, mechanical, chubby, sunken, expressive, wrinkled, bearded / with moustache, or with make-up), and the presence of a mouth (e.g., big, small, squared, smiling, minimalistic, moveable, red, featuring full lips, or featuring thin lips)) have a positive direct effect on active facilitation, passive harm and passive facilitation and a positive indirect effect on active facilitation, passive harm and passive facilitation mediated by pity, admiration, communion and femininity.

4. Methodology

In the following, I will explain how the data in the study was collected. I will detail how the set of 70 anthropomorphic robots were selected, along with how participants were recruited for this research, and will describe the procedure and measures of this study.

4.1 Study design

In order to understand which appearance qualities affect emotional reactions and behavioural tendencies, an online survey on LimeSurvey was conducted. A set of 70 robots from the Anthropomorphic roBOTS (ABOT) dataset (Philips et al., 2018) were selected, and based on Mieczkowski and colleagues (2019), I asked at least 10 participants to evaluate each image of the robot, leading to a total of 700 participants required for this research. Participants were randomly allocated to one of the images, and were asked to rate the robot based on their perceptions of the robot's appearance, social categorization (in terms of gender and age), stereotypical traits (in terms of communion and agency), emotional reactions (in terms of admiration, pity, envy and contempt), and behavioural tendencies (in terms of active facilitation, passive facilitation, active harm and passive harm).

4.2 Robot selection

The 70 robots used in this study were sampled in two rounds: Initially, I started to select robots from Perugia and colleagues' (2023) 80 robot images from the ABOT dataset. I opted to use this dataset because it offered a classification of robots into four distinct clusters, such as cluster 1 representing robots with low agency and low communion scores. To ensure representation from each cluster, the top five robots with the highest agency scores, the top five with the highest communion scores, as well as the bottom five with the lowest agency scores and the bottom five with the lowest communion scores were selected. This approach guaranteed that the sample of robots would reflect the respective scores on these stereotypical dimensions. Moreover, including robots with varying scores of agency and communion was crucial for this research, to identify prominent appearance qualities and explore potential associations between different agency and communion scores and these qualities. Since some clusters exhibited univalent stereotypes (e.g.,

low agency/ low communion), a total of 52 robots were selected in the first round, accounting for instances where a robot possessed both low agency and low communion scores.

In the second round, it became clear that a majority of the 52 robots collected in the first step had higher scores on masculinity than femininity. Given I wanted to have an equal representation of male and female robots, an additional 18 robots that had high scores on femininity in the ROBOts - Gender and Age Perception (ROBO-GAP) dataset (Perugia et al., 2022) were selected. This equal representation and including more feminine robots into this research was important, since previous research (Perugia et al., 2022) worked with a predominantly masculine sample and had difficulty discriminating whether the lack of significant effects of femininity on stereotyping were valid findings or due to the limited number of female robots with respect to male ones. I did not include additional robots with high scores on gender neutrality as gender neutrality is not known to affect stereotypical traits of communion and agency (Perugia et al., 2023). Of the 70 robots in this study, 32 of the robots were predominantly masculine (46%), 32 of the robots were predominantly feminine (46%), and 6 of the robots were predominantly gender neutral (8%) based on the ratings of the ROBO-GAP dataset (Perugia et al., 2022). Overall, the robots were perceived to be 29.7 years old on average. Fourteen were perceived to be younger than 20 years old (20%), 24 were perceived to be between 20 to 30 years old (34.29%), 29 were perceived to be between 30 to 50 years old (41.43%), 2 were perceived to be between 50 and 75 years old (2.86%), and 1 robot (i.e., Albert Einstein Hubo) was perceived to be over 75 years old (1.43%). The images of the robots can be found in the appendix (Appendix D).

4.3 Participants

714 participants were recruited to participate in the study using the Prolific platform. Participants in the survey had different nationalities (50 in total), the majority being South Africa (151 participants, 21.70%), Portugal (115 participants, 16.52%), Poland (78 participants, 11.21%) and the United Kingdom (70 participants, 10.06%) (see Table 1, and Appendix A for the full overview). Participation in this study was voluntary, and participants were compensated with £1.50 upon completion of the survey. The study was approved by the Ethical Review Board of the Human-Technology Interaction department at the Eindhoven University of Technology.

Table 1*Demographic overview of participants in the study.*

Women	Men	Non-binary	Prefer not to say	Age				Nationality (N)
				<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	
326	295	90	3	28.78	9.19	18	76	50

Note: For women and men, I considered both cis-gender and trans-gender participants.

4.4 Procedure and Measures

Before the actual survey began, participants read a description of the research and were asked to give their consent. After consent was provided, each participant was randomly shown one of the 70 images of robots, which they continued to see throughout the whole survey. All robots were rated by at least 10 participants with a maximum of 13 responses (e.g., for the robot Asimo).

Participants first saw the questions related to the attribution of social categories, stereotypical traits, emotional reactions, and behavioural tendencies. Then, they were asked to evaluate the presence of certain appearance qualities in the robot. To ensure that participants did not determine the true purpose of the study, the order of the presentation of the questions regarding social categories, stereotypical traits, emotional reactions, and behavioural tendencies was randomised. However, to make rating the robot's appearance qualities easier and faster for participants, the questions regarding the appearance qualities of the robot were asked starting from the top of the robot's body (e.g., the head) to the bottom (e.g., the legs). The questions on the appearance qualities were not randomised to ensure that participants could look at the robot once from head to toe without having to look at different body parts of the robot each time. The images of the robots were kept in their original size in the ABOT database.

Overall, participants were asked to answer a total of 63 questions. The survey took on average around 6 minutes to complete. The first set of questions asked about the robot's perceived age: *In which moment of life would you place the robot in the image considering that 1 is like a newborn child and 100 like a very old person?* (Perugia et al., 2022). Participants had to indicate the robot's age by moving a slider. Furthermore, participants were asked to evaluate the perceived gender of the robot: *How would you describe the robot in the image?* Participants could rate their

perception on three items *feminine*, *masculine*, and *gender neutral*, using a 7-point Likert scale (1= completely disagree, 7= completely agree) (Perugia et al., 2022). As for stereotypical traits, similar to Perugia and colleagues (2023), participants were presented with statements on communion, and asked to rate to what extent they agreed with them: *This robot is [affectionate, compassionate, tender, gentle, sympathetic]*, using a 7-point Likert scale (1= strongly disagree, 7= strongly agree; Short version of the Bem Sex Role Inventory by Choi et al., 2009). Similarly, participants were asked about the perceived agency of the robot, by rating statements such as: *This robot is [able to defend its own beliefs, willing to take a stand, has leadership abilities, able to make decisions easily, has a strong personality]*, using a 7-point Likert scale (1= strongly disagree, 7= strongly agree) (Short version of the Bem Sex Role Inventory by Choi et al., 2009). Additionally, participants were asked to rate questions regarding the emotional reactions that the robot evoked: *To what extent do you feel [emotion, e.g. envy] towards this robot?*, using a 7-point Likert scale (1= not at all, 7=extremely), with two items for each emotional reaction: envy (envy, jealousy), admiration (admiration, pride), pity (pity, sympathy) and contempt (contempt, disgust) (Mieczkowski et al., 2019). Moreover, participants were asked to rate questions regarding the behavioural tendencies that the robot evoked: *How likely would you be to [behaviour, e.g. help] this robot?*, using a 7-point Likert scale (1= not at all, 7= extremely), with two items for each behavioural tendency: active facilitation (help, protect), active harm (fight, attack), passive facilitation (cooperate with, associate with) and passive harm (exclude, demean) (Mieczkowski et al., 2019). While Cuddy and colleagues (2007) asked questions about social groups and common perceptions of them, I asked the question about specific robots and each participant's perception of them as done by Mieczkowski and colleagues (2019) and Perugia and colleagues (2022 and 2023), since robots are not widely used in society, global views of robots' social groups are not available yet. The full questionnaire used in the study can be found in the appendix (Appendix B).

To identify the appearance qualities that participants should rate, I started with the appearance cues identified by Philips and colleagues (2018), which included *surface look* features – presence of eyelashes, head hair, skin, nose, eyebrows, and apparel, *body manipulators* – presence of arms, torso, fingers, and legs, and *facial features* – presence of face, eyes, head, and mouth of robots. For each of the overarching appearance cues (e.g., surface look) I asked questions regarding the qualities of the cues (Appendix B). Hence, instead of looking at the presence or absence of head hair, I also looked at if the robot had *short, long, metallic, human-like, blond, dark*

etc. hair, or instead of looking at the presence of absence of eyebrows, I asked participants to evaluate whether the eyebrows were *thick, thin, metallic, human-like, straight, curved, frowning* etc. I first asked participants to indicate whether the cue was present in the image of the robot or not: *Does the robot have visible skin?* (answer options: *yes, no*), and if the answer was *yes*, another question appeared where the participant could select different qualities of the cue: *Is the skin: metallic, realistic (life-like, human-like), black, brown, white, pink, yellow or other colour [blank field]* (Figure 3).

Figure 3

LimeSurvey questions regarding the qualities of appearance qualities of the robot.

Perception of a Robot
Please, observe the robot in the image carefully.

Please select the **option** that applies when looking at the robot in the image.
*

Only answer this question if the following conditions are met:
Answer was 'I consent to take part in this study' at question '1 [Participate]' (Certificate of consent (participation in the study) By starting this study, I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.)

Please choose the appropriate response for each item:

	Yes	No
Does the robot have visible skin?	<input type="radio"/>	<input type="radio"/>

Is the skin: *

Only answer this question if the following conditions are met:
Answer was 'Yes' at question '36 [Skin?]' (Please select the option that applies when looking at the robot in the image. (Does the robot have visible skin?))

Check all that apply
Please choose all that apply:

Metallic
 Realistic (life-like, human-like)
 Black
 Brown
 White
 Pink
 Yellow
 Other color:

Note: In this example participants were asked about the qualities of the ‘skin’ appearance cue.

After evaluating the robot, participants were asked an additional set of 18 questions regarding their own demographic information and individual characteristics. First participants were asked questions regarding their gender, by ticking all boxes that applied to them: [*woman, man, non-binary, I'd rather not say, I'd rather specify specify (blank field)*] and age, by stating their age in numbers. Next, individual characteristics, such as their familiarity with Artificial Intelligence (AI), robots, and sci-fi were asked, *Rate your degree of familiarity with... [AI, Robots, and Science fiction]* using a 7-point Likert scale (1= not at all, 7= very familiar) (Perugia et al., 2022). Furthermore, participants were asked to state their anthropomorphism tendencies, *please, rate the extent to which you agree with the following statements [e.g. To what extent does the average robot have consciousness?]*, using a 11-point Likert scale (0= not at all, 10= very much) (Waytz et al., 2010). Finally, participants answered questions regarding their (sub)conscious

benevolent and hostile sexism, *please, rate the extent to which you agree with the following statements [e.g. Many women have a quality of purity that few men possess (B), or women exaggerate problems they have at work (H)]*, using a 7-point Likert scale (1= strongly disagree, 7= strongly agree; B= Benevolent Sexism; H= Hostile Sexism) (Ambivalent Sexism Inventory – Short version; Rollero et al., 2014) (Appendix B). Some demographics (i.e., the nationality of participants) were provided by Prolific. The order of these questions were not randomised, as it concerned only the individual characteristics of participants where there was no concern about an effect in which the questions were presented.

5. Results

5.1 Data screening

Before running the actual analyses, I wanted to ensure that only the responses of participants who filled out the survey intentionally were included. A total of 740 participants' responses were collected, from which the participants who could not see the image of the robot were removed first ($N = 6$). I then went on to check how long it took participants to fill out the survey (on average this was 6 minutes), and excluded those that completed the survey in 5 minutes or less, as these participants did not take enough time to read and understand the questions properly ($N = 9$). Furthermore, participants who gave the same answer to all questions on communion, agency, emotional reactions, and behavioural tendencies were removed ($N = 2$). Similarly, the participants who gave the same scores on their individual traits such as familiarity with AI, robots, and sci-fi, their anthropomorphism tendencies as well as their benevolent and hostile sexism ratings were removed ($N = 9$). In total, I worked with a data set of 714 participants in the following analyses.

5.2 Effect of appearance qualities on emotional reactions (RQ1) and behavioural tendencies (RQ2)

This research aims to understand the effects that appearance qualities have on emotional reactions (RQ1) and behavioural tendencies (RQ2). Before answering the research questions, I ran two pre-analyses: (1) to determine which appearance qualities should be used in the analysis, and (2) to understand whether social categories (in terms of gender and age) and stereotypical traits (in terms of communion and agency) are potential mediators of any direct effects that appearance qualities might have. In addition, for RQ2, I discovered by running the analysis on RQ1 whether emotional reactions (i.e., pity, envy, admiration and contempt) could be additional potential mediators. For the first pre-analysis step (1), I ran a multiple linear regression to determine which of the theory-based appearance qualities were predictors of social categories. The appearance qualities that were found to have a significant effect on masculinity, femininity and age were the ones I selected to work with in the remaining analysis. In the second pre-analysis step (2), I then ran a multiple linear regression of these appearance qualities on stereotypical traits communion and agency. Here, I was able to determine whether communion or agency could be potential mediators of any effects of appearance qualities. After the pre-analyses were completed, a multiple linear regression of appearance qualities on emotional reactions was run (RQ1), followed by a subsequent multiple

linear regression with appearance qualities on behavioural tendencies (RQ2) to answer both research questions. In these regressions the direct effects of appearance qualities on emotional reactions and behavioural tendencies was determined. I followed up these regressions by running separate mediation analyses to determine whether the direct effects of the qualities on emotional reactions and behavioural tendencies remained, or whether they were mediated by social categories, stereotypical traits and (for RQ2) by emotional reactions. In the following, I will explain each analysis and outcome in detail. For the sake of clarity and brevity, in this report, I present only the findings that were significant. The non-significant findings, of the appearance qualities that were not found to have a significant effect on social categories, can be found in the appendix (Appendix C).

5.2.1 Pre-analysis: Selection of appearance qualities

Before being able to test research questions RQ1 and RQ2, I wanted to determine which specific appearance qualities, found in the previous literature review, were significant predictors of the social categories gender and age in this study. These would be the appearance qualities that would be worked with consistently in the subsequent analyses. In total, three multiple linear regressions were run; 1. appearance qualities on masculinity, 2. appearance qualities on femininity, and 3. appearance qualities on age.

In the first model, I tested which appearance qualities had an effect on masculinity. For this multiple linear regression model normality was rejected, which is why a nonparametric regression was run instead. The overall model was significant, $R^2 = 0.123$, $F(22, 691) = 5.9$, $p < 0.0001$. The appearance qualities of muscular torso, chubby torso, torso with visible mechanics, and legs were found to be significant positive predictors of masculinity, whereas the appearance qualities of slender torso was a significant negative predictor of masculinity (Table 2).

Table 2*Appearance qualities that have an effect on masculinity.*

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Muscular torso	0.829	0.256	3.24	0.124	0.001
Chubby torso	0.573	0.264	2.17	0.081	0.030
Torso with mechanics visible	0.431	0.173	2.49	0.108	0.013
The presence of legs	0.785	0.298	2.63	0.207	0.009
Slender torso	-0.467	0.198	-2.35	-0.089	0.019

Note: Regression model: $R^2 = 0.123$, $F(22, 691) = 5.9$, $p < 0.0001$.

In the second model, the effect of appearance qualities on femininity was examined. For this multiple linear regression model normality was rejected, which is why a nonparametric regression was run instead. Overall, this model was found to be significant, $R^2 = 0.283$, $F(24, 690) = 16$, $p < 0.0001$. The appearance qualities of skirt, dress, pink apparel, face with make-up, and the presence of eyelashes were significant positive predictors of femininity, whereas blue apparel and dark apparel were found to be significant negative predictors of femininity (see Table 3).

Table 3*Appearance qualities that have an effect on femininity.*

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Skirt	1.162	0.320	3.63	0.126	0.0001
Dress	1.031	0.263	3.92	0.172	0.0001
Pink apparel	1.775	0.454	3.91	0.117	0.0001
Face with make-up	1.225	0.223	5.48	0.200	0.0001
The presence of eyelashes	1.008	0.244	4.13	0.174	0.0001
Blue apparel	-1.181	0.239	-4.94	-0.149	0.0001
Dark apparel	-0.623	0.270	-2.30	-0.075	0.022

Note: Regression model: $R^2 = 0.283$, $F(24, 690) = 16$, $p < 0.0001$.

In the last model, the effect of appearance qualities on age was tested. For this multiple linear regression model normality was rejected, which is why a nonparametric regression was run instead. The overall model was found to be significant, $R^2 = 0.115$, $F(10, 703) = 9.98$, $p < 0.0001$. The appearance qualities of the presence of a nose, grey hair and white hair were found to be significant positive predictors of age, whereas big eyes and a short nose were negative predictors of age (Table 4).

Table 4*Appearance qualities that have an effect on age.*

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
The presence of a nose	10.556	2.67	3.94	0.253	0.0001
Grey hair	24.584	6.27	3.92	0.152	0.0001
White hair	15.147	4.50	3.37	0.125	0.0001
Big eyes	-3.510	1.50	-2.33	-0.082	0.020
Short nose	-8.272	2.25	-3.67	-0.145	0.0001

Note: Regression model: $R^2 = 0.115$, $F(10, 703) = 9.98$, $p < 0.0001$.

In this first pre-analysis, I found which appearance qualities had a significant effect on masculinity, femininity and age. While Perugia and colleagues (2022) I found that the presence of a torso predicted masculinity, here it was found that rather the quality of it being muscular, chubby and having visible mechanics are significant positive predictors of masculinity (*H1b*). Furthermore, the presence of legs (*H1f*) was a significant positive predictor of masculinity as discovered by Perugia and colleagues (2022). Additionally, I found that a slender torso was a negative predictor of masculinity (*H1b*). For femininity, not only the presence of apparel as found by Perugia and colleagues (2022) was a predictor of femininity, but rather specific clothing items such as a skirt and a dress (*H1j*). Items that were pink were associated with femininity, whereas darker tones of clothing (i.e., blue and dark apparel) were significant negative predictors of femininity (*H1j*). Furthermore, I found that a face with make-up (*H1h*) as well as the presence of eyelashes (*H1k*) were significant positive predictors of femininity. These findings are in-line with previous work by Eyssel and Hegel (2012). Finally for age I found that, contrary to Kalegina and colleagues (2018), the presence of hair did not have a negative effect on age, but rather the quality of it being white and grey were positive predictors of age (*H1p*). Additionally, findings differed from Kalegina and colleagues (2018), as they found the presence of a nose to be a negative predictor of age, which in this research was found to be a positive one (*H1o*). I did, however, find that the quality of a short nose was a negative predictor of age (*H1o*). In-line with Kalegina and colleagues (2018), big eyes negatively predicted age (*H1n*). Listed in figure 4 below are the appearance qualities that were used in the subsequent analysis.

Figure 4

Appearance qualities that have a significant effect on social categories masculinity, femininity and age.



5.2.2 Pre-analysis: Appearance qualities on stereotypical traits

After selecting the appearance qualities, I went on to determine the effect that these qualities had on the stereotypical traits communion and agency. Knowing which appearance qualities predict stereotypical traits is important for later analyses when studying the effects of appearance qualities on emotional reactions (RQ1) and behavioural tendencies (RQ2), to determine whether communion or agency are potential mediators of any direct effects. In order to know whether there was a direct effect of the appearance qualities on communion and agency, multiple linear regressions with the appearance qualities as predictors on either communion or agency were run, and then I checked through individual mediation analyses how much of this direct effect was mediated by social categories. In the following, the figures will show the mediations that were examined.

In the first model, I tested the effect of appearance qualities (that predict masculinity) on agency (H2a). In a multiple linear regression analysis the effect of these qualities on agency was tested, and for this multiple linear regression model normality was rejected, which is why a

nonparametric regression was run instead. I found that the overall model was significant, $R^2 = 0.043$, $F(6, 697) = 5.16$, $p < 0.0001$. The appearance qualities that significantly predicted agency were a muscular torso, a chubby torso and a torso with visible mechanics (Table 5).

Table 5

Appearance qualities (that predict masculinity) that have an effect on agency.

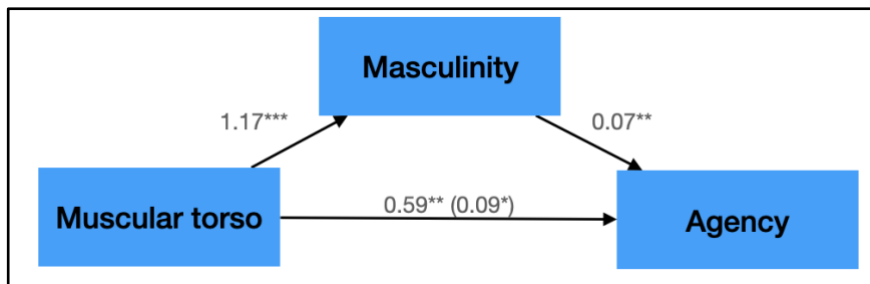
Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (B)	Significance level (p)
Muscular torso	0.681	0.180	3.78	0.144	0.0001
Chubby torso	-0.528	0.180	-2.93	-0.105	0.004
Torso with mechanics visible	-0.265	0.110	-2.41	-0.094	0.016

Note: Regression model: $R^2 = 0.043$, $F(6, 697) = 5.16$, $p < 0.0001$.

I then went on to check whether masculinity predicted agency, as to say, whether masculinity could mediate the direct effect of the appearance qualities on agency, and found that masculinity was a significant predictor of agency, $b = 0.09$, $t(704) = 3.53$, $p < 0.0001$. To determine how much of the direct effects of appearance qualities on agency were mediated by masculinity, I ran three separate mediation analyses using the appearance qualities that predict masculinity (Figure 5) as predictors of agency, and masculinity as a potential mediator. For the effect of muscular torso I found that the direct effect of the quality on agency remained, and that there was a significant mediation of muscular torso via masculinity on agency which did not change the direct effect of the qualities itself (Figure 5). In the following mediations I report the unstandardized coefficients, as this is in-line with mediations reported by Cuddy and colleagues (2007) and Mieczkowski and colleagues (2019).

Figure 5

The effect of muscular torso on agency mediated by masculinity.

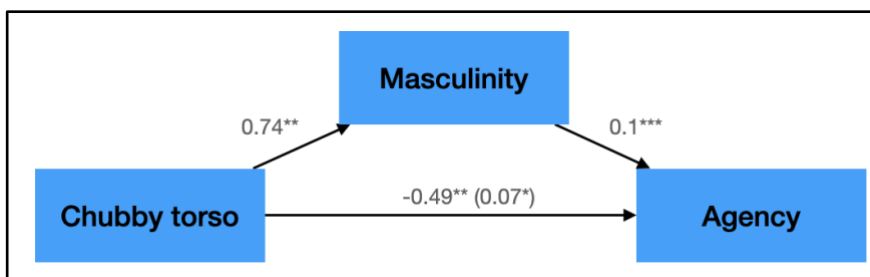


Note: Unstandardized coefficients are reported, with the respected p-values * $p < .05$., ** $p < .01$., *** $p < .001$. Total effects are within the parentheses, and direct effects are adjacent. Full line means significant effect, a dotted line means no significant effect.

In the second mediation, I found that the direct effect of chubby torso on agency remained. Again, there was a significant mediation of chubby torso via masculinity on agency, but this effect did not change the direct effect of the quality on agency (Figure 6).

Figure 6

The effect of chubby torso on agency mediated by masculinity.

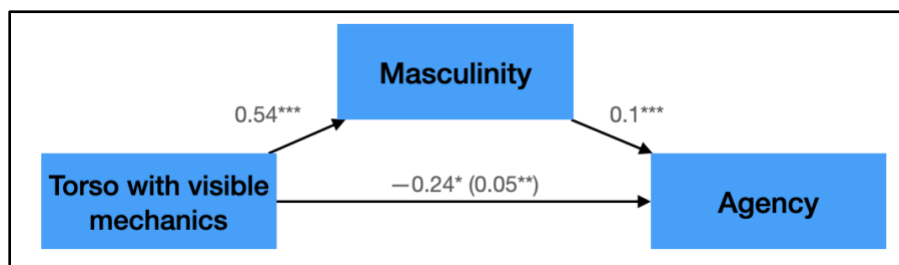


Note: Unstandardized coefficients are reported, with the respected p-values * $p < .05$., ** $p < .01$., *** $p < .001$. Total effects are within the parentheses, and direct effects are adjacent. Full line means significant effect, a dotted line means no significant effect.

Finally, it was discovered that the effect of a torso with mechanics visible on agency remained, and again there was a significant mediation of torso with mechanics visible via masculinity on agency, however, this did not change the direct effect that torso with mechanics visible had on agency (Figure 7).

Figure 7

The effect of chubby torso on agency mediated by masculinity.



Note: Unstandardized coefficients are reported, with the respected p-values *p <.05., **p <.01., ***p<.001. Total effects are within the parentheses, and direct effects are adjacent. Full line means significant effect, a dotted line means no significant effect.

In the second model, the effect of appearance qualities (that predict femininity) on communion was tested. In a multiple linear regression analysis, the effect of these qualities on communion were tested, and for this multiple linear regression model normality was rejected, which is why a nonparametric regression was run instead. I found that the overall model was not significant, $R^2 = 0.0197$, $F(7, 705) = 1.96$, $p < 0.0582$. Of all the appearance qualities that significantly predicted femininity, only pink apparel ($b = 0.112$, $t(705) = 2.84$, $p < 0.005$) had a significant positive effect on communion (Table 6).

Table 6

Appearance qualities (that predict femininity) that have an effect on communion.

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (B)	Significance level (p)
Pink apparel	1.293	0.455	2.84	0.112	0.005

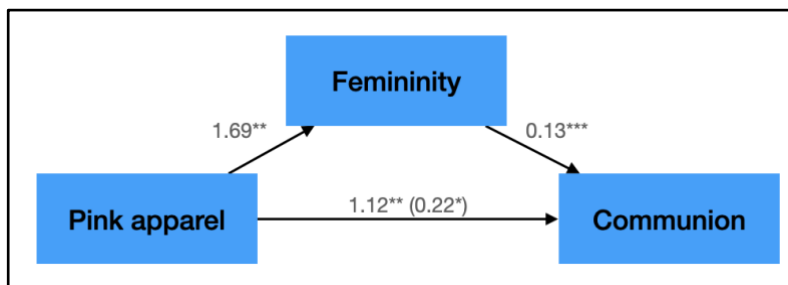
Note: Regression model: $R^2 = 0.0197$, $F(7, 705) = 1.96$, $p < 0.0582$.

Then I went on to check whether femininity and communion were correlated, as to say, whether femininity could be a potential mediator of the direct effect of pink apparel on communion, and found that indeed femininity was a significant predictor of communion, $R^2 = 0.033$, $F(1, 711) = 24.34$, $p < 0.00001$. Subsequently, I determined whether the effect of pink apparel on communion would vanish when mediated by femininity, and I discovered that the direct effect of pink apparel remained, and that, while there was a significant mediation of pink apparel

via femininity on communion, this did not change the direct effect that pink apparel had on communion (Figure 8).

Figure 8

The effect of pink apparel on communion mediated by femininity.



Note: Unstandardized coefficients are reported, with the respected p-values *p <.05., **p <.01., ***p<.001. Total effects are within the parentheses, and direct effects are adjacent. Full line means significant effect, a dotted line means no significant effect.

In the third model, the effect of appearance qualities (that predict age) on communion were tested. In a multiple linear regression analysis the effect of these qualities on communion were tested, and for this multiple linear regression model normality was rejected, which is why a nonparametric regression was run instead. I found that the overall model was significant, $R^2 = 0.0192$, $F(5, 707) = 3.32$, $p < 0.0057$. Of all the appearance qualities that significantly predicted age, grey hair ($b = 0.106$, $t(713) = 3.14$, $p < 0.002$) was the only significant predictor of communion (see Table 7). To determine whether this effect was direct or whether age mediated it, I went on to determine the effect of age on communion and found that age does not significantly predict communion ($R^2 = 0.0013$, $F(1, 711) = 0.95$, $p < 0.3309$). Hence, age could not act as a mediator and the effect of grey hair on communion remained to have a direct effect.

Table 7

Appearance qualities (that predict age) that have an effect on communion.

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Grey hair	1.277	0.406	3.14	0.106	0.002

Note: Regression model: $R^2 = 0.0192$, $F(5, 707) = 3.32$, $p < 0.0057$.

In this second pre-analysis, muscular torso was found to have a positive effect on agency (*H2a*), chubby torso and torso with visible mechanics had a negative effect on agency, pink apparel had a positive effect on communion (*H2b*), and finally, grey hair had a positive effect on communion (*H2c*). None of the direct effects of appearance qualities on communion or agency were fully mediated by social categories. However, all mediations were significant, for instance the mediated effect of muscular torso via masculinity on agency was significant (*H2a*). Nevertheless, the direct effects of the appearance qualities remained significant too despite any mediations. The appearance qualities (e.g., muscular torso) that predict masculinity had through masculinity had a positive indirect effect on agency. The same results were found for the appearance qualities that predicted femininity, which was that pink apparel had a positive effect on femininity, which had an indirect positive effect on communion (*H2b*). For qualities that predicted age, however, no significant indirect effects were found, which is why only partial proof of *H2c* was found (i.e., grey hair has a direct positive effect on communion). Furthermore, the direct effect of the appearance qualities on social categories were not always aligned with the direction of the mediated effects. For instance, the direct effect of torso with visible mechanics on agency was negative, whereas the mediated effect of torso with visible mechanics via masculinity on agency was positive (Figure 7). Similarly, a negative direct effect of chubby torso on agency was found, whereas a positive effect of chubby torso via masculinity on agency, which is why *H2a* was only partially accepted (Figure 6).

5.2.3 RQ1: Appearance qualities on emotional reactions

In the previous section I learned which appearance qualities are predictors of stereotypical traits communion and agency. With this knowledge, I was then able to test the first research question by running in a multiple linear regression the direct effect of appearance qualities on emotional reactions (RQ1), and following up on these effects by conducting individual mediation analyses for each appearance quality, with social categories (in terms of gender and age) and stereotypical traits (communion and agency) as mediators, to determine which direct effects remained. In the following, normality for all regressions was rejected, which is why nonparametric results are reported.

In a multiple linear regression, the effects of appearance qualities (that predict masculinity) on emotional reactions were found. The first significant effects of appearance qualities were found

on envy. While the overall model was not significant ($R^2 = 0.0097$, $F(6, 707) = 1.18$, $p < 0.3150$), torso with visible mechanics was a significant negative predictor of envy (Table 8).

Table 8

Appearance qualities (that predict masculinity) that have an effect on envy.

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Torso with visible mechanics	-0.139	0.063	-2.18	-0.083	0.029

Note: Regression model: $R^2 = 0.0097$, $F(6, 707) = 1.18$, $p < 0.3150$.

Furthermore, in the model with appearance qualities on admiration ($R^2 = 0.033$, $F(6, 707) = 3.20$, $p < 0.0042$), there was a significant positive effect of muscular torso on admiration (see Table 9).

Table 9

Appearance qualities (that predict masculinity) that have an effect on admiration.

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Muscular torso	0.792	0.215	3.68	0.161	0.0001

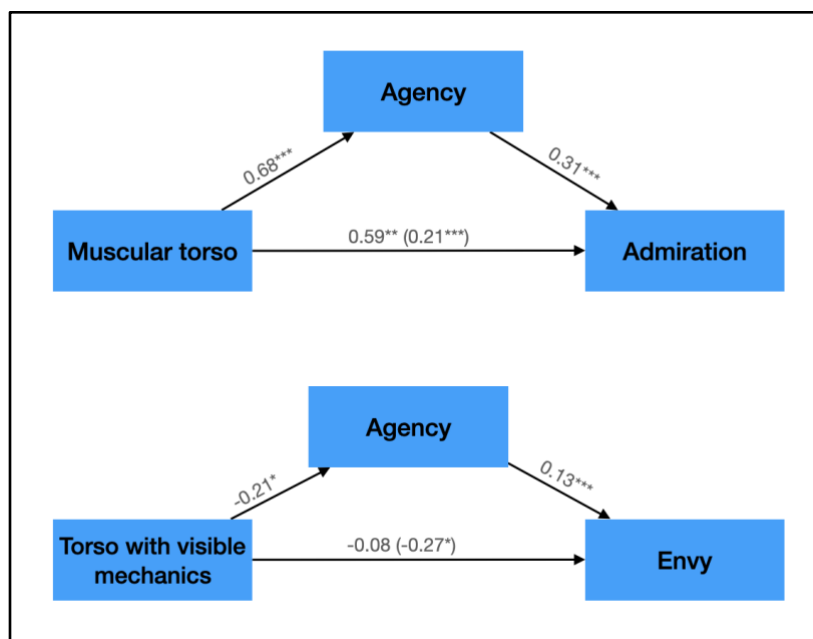
Note: Regression model: $R^2 = 0.033$, $F(6, 707) = 3.20$, $p < 0.0042$.

In bivariate correlations, I then went on to check whether masculinity and agency were correlated with the four emotional reactions, as to say, whether masculinity and agency could be potential mediators of the direct effect of appearance qualities on emotional reactions. While masculinity was not correlated to any of the four emotions, agency was found to be a significant predictor of contempt ($R^2 = 0.0171$, $F(1, 702) = 12.14$, $p < 0.0005$), envy ($R^2 = 0.055$, $F(1, 702) = 36.31$, $p < 0.0001$), pity ($R^2 = 0.041$, $F(1, 702) = 32.47$, $p < 0.0001$) and admiration ($R^2 = 0.106$, $F(1, 702) = 75.70$, $p < 0.0001$). Subsequently, I went on to determine how much of the above effects of appearance qualities on emotional reactions were mediated by social categories and stereotypical traits, and found that the direct effect of muscular torso on admiration remained, and that there was a significant mediation of muscular torso via agency on admiration, which did not change the direct effect of muscular torso on admiration (Figure 9). I ran the same mediation this time for the effect of torso with visible mechanics on envy, but the effect of this quality vanished

as there was a significant mediation of torso with mechanics via agency on envy. In this model, the appearance qualities of a muscular torso had a direct, non-mediated effect on admiration.

Figure 9

The effect of appearance qualities that affect masculinity on emotional reactions mediated by agency.



Note: Unstandardized coefficients are reported, with the respective p-values * $p < .05$., ** $p < .01$., *** $p < .001$. Total effects are within the parentheses, and direct effects are adjacent. Full line means direct effect, a dotted line means no direct effect.

In the second model, I tested the effect of appearance qualities (that predict femininity) on emotional reactions. In a multiple linear regression analysis I found an overall significant regression model on envy ($R^2 = 0.014$, $F(7, 706) = 2.74$, $p < 0.0081$). Blue apparel and dark apparel had a significant negative direct effect on envy (Table 10).

Table 10*Appearance qualities (that predict femininity) that have an effect on envy.*

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Blue apparel	-0.221	0.095	-2.34	-0.071	0.020
Dark apparel	-0.186	0.072	-2.59	-0.057	0.010

Note: Regression model: $R^2 = 0.014$, $F(7, 706) = 2.74$, $p < 0.0081$.

Furthermore, a multiple linear regression with appearance qualities on admiration was run ($R^2 = 0.0103$, $F(7, 706) = 1.25$, $p < 0.2734$) and found that face with make-up had a significant negative effect (see Table 11).

Table 11*Appearance qualities (that predict femininity) that have an effect on admiration.*

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Face with make-up	-0.391	0.152	-2.57	-0.092	0.010

Note: Regression model: $R^2 = 0.0103$, $F(7, 706) = 1.25$, $p < 0.2734$.

Here, I did not have to check for any mediation effects of femininity or communion, as femininity was not a significant predictor of any of the four emotions. Similarly, communion could not be a mediator either as face with make-up was not a significant predictor of communion in the first place. Hence, I found a significant negative effect of blue apparel and dark apparel on envy, and a significant negative effect of face with make-up on admiration.

In the last model, I tested the effects of appearance qualities (that predict age) on emotional reactions. In a multiple linear regression analysis I found an overall significant regression model on contempt ($R^2 = 0.038$, $F(5, 708) = 5.05$, $p < 0.0001$). Big eyes, the presence of a nose and white hair were significant predictors of contempt (Table 12).

Table 12*Appearance qualities (that predict age) that have an effect on contempt.*

Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Big eyes	0.194	0.086	2.24	0.083	0.026
The presence of a nose	0.404	0.100	4.03	0.176	0.0001
White hair	0.458	0.232	1.97	0.069	0.049

Note: Regression model: $R^2 = 0.038$, $F(5, 708) = 5.05$, $p < 0.0001$.

Furthermore, I found a significant positive effect of big eyes on pity (see Table 13), where the overall model was also found to be significant ($R^2 = 0.021$, $F(5, 708) = 3.14$, $p < 0.0082$).

Table 13*Appearance qualities (that predict age) that have an effect on pity.*

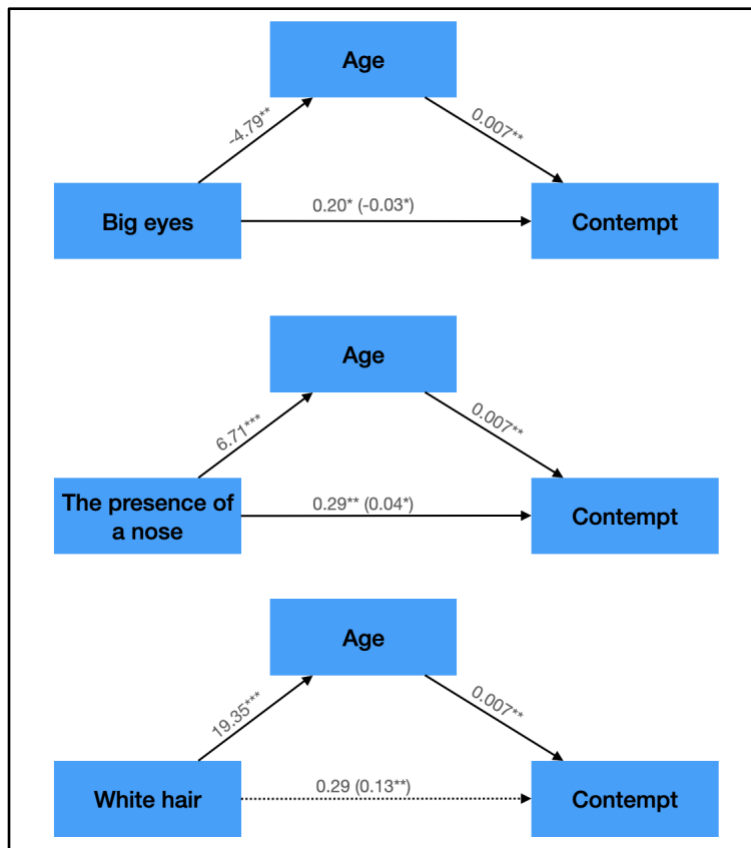
Appearance qualities	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Big eyes	0.349	0.096	3.63	0.136	0.0001

Note: Regression model: $R^2 = 0.021$, $F(5, 708) = 3.14$, $p < 0.0082$.

I then went on to determine how much of the above effects of appearance qualities on emotional reactions were mediated by age (Figure 10).

Figure 10

The effect of appearance qualities that affect age on emotional reactions mediated by age.



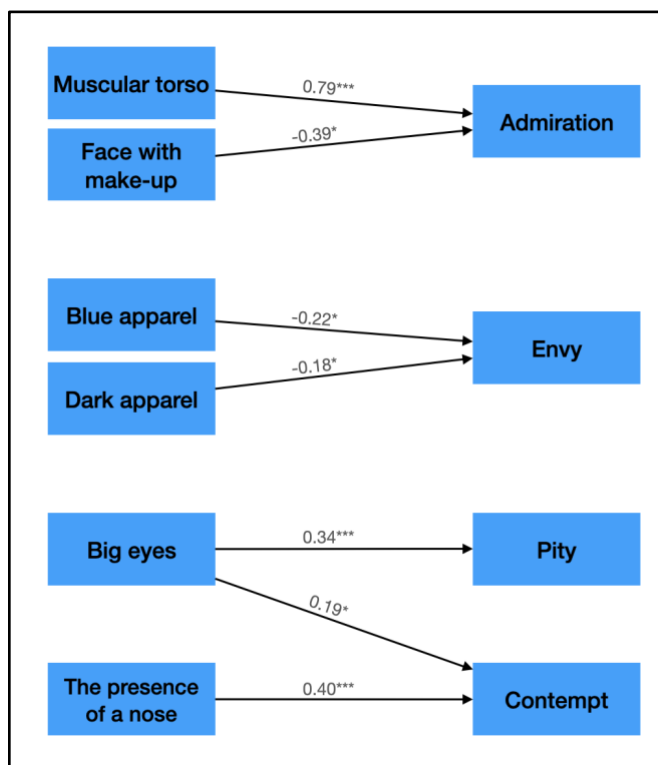
Note: Unstandardized coefficients are reported, with the respected p-values * $p < .05$., ** $p < .01$., *** $p < .001$. Total effects are within the parentheses, and direct effects are adjacent. Full line means direct effect, a dotted line means no direct effect. Only significant mediations are visualised here.

While a significant mediation of big eyes via age on contempt was found, the direct effect of big eyes on contempt remained. I then checked how much of the effect of big eyes on pity was mediated by communion, as communion was a significant predictor of pity. I found that the direct effect of big eyes on pity remained, and that communion was not a significant mediator of the effect of big eyes on pity. Similarly, the effect of the presence of a nose on contempt remained even though age was a significant mediator of the effect of presence of a nose via age on contempt. Lastly, the effect of white hair on contempt vanished, and age acted as a significant mediator of the effect of white hair via age on contempt. In total, there were direct effects of big eyes on contempt and pity, and the presence of a nose on contempt.

Overall, for RQ1 I found that a muscular torso is a positive predictor of admiration (*H3a*), blue and dark apparel are negative predictors of envy, face with make-up is a negative predictor of admiration, big eyes is a positive predictor of contempt and pity (*H3c*), and the presence of a nose is a positive predictor of contempt (Figure 11). What became noticeable in these analyses is that there were only two mediators, age and agency. Communion, femininity and masculinity were not found to be significant predictors of emotional reactions. Additionally, most of the effect of the appearance qualities on emotional reactions were direct and not mediated, only the effects of torso with mechanics visible via agency on envy and white hair via age on contempt were mediated. However, these were not full mediations as for torso with visible mechanics there was no mediation by masculinity, and for white hair on contempt there was no mediation of communion, which would have made it a complete mediation (as predicted in *H3a* and *H3c*).

Figure 11

Appearance qualities that have a significant effect on emotional reactions.



Note: Unstandardized coefficients of the regression analyses are reported, with the respected p-values * $p < .05$., ** $p < .01$., *** $p < .001$.

5.2.4 RQ2: Appearance qualities on behavioural tendencies

Up until now I have examined the first research question (RQ1), which was aimed at determining which appearance qualities affect emotional reactions. I now continue to uncover the subsequent research question (RQ2) which is aimed at determining the link between appearance qualities and behavioural tendencies. This question was tested in a similar way as RQ1, by first running multiple linear regressions using appearance qualities as predictors of behavioural tendencies, and then checking, in separate mediation analyses, whether the direct effects of appearance qualities remained when mediated by social categories and stereotypical traits. In these mediation analyses emotional reactions (i.e., contempt, envy, pity, and admiration) were also included as additional potential mediators. In the following, normality for all regressions was rejected, which is why nonparametric results are reported.

In the first model, the effect of appearance qualities (that predict masculinity) on behavioural tendencies were tested. In a multiple linear regression analysis I tested the effect of these qualities on behavioural tendencies. The first significant effects of appearance qualities was found on active facilitation, where the overall model was significant ($R^2 = 0.030$, $F(6, 690) = 4.04$, $p < 0.0005$), and muscular torso was a significant predictor of active facilitation (Table 14). Furthermore, I found the chubby torso to be a significant negative predictor of active harm, where the overall model of appearance qualities on active harm was not significant ($R^2 = 0.011$, $F(6, 689) = 1.70$, $p < 0.1182$) (Table 15). Additionally, I found that muscular torso was found to be a significant positive predictor of passive facilitation, in a regression model that was significant too ($R^2 = 0.017$, $F(6, 692) = 2.14$, $p < 0.0472$) (Table 16).

I then went on to check whether masculinity, agency, and the four emotional reactions were potential predictors of the four behavioural tendencies, and thus, whether these could mediate any direct effects of appearance qualities on behavioural tendencies. Masculinity was found to be a significant predictor of passive harm (Table 17). Based on previous research from Cuddy and colleagues (2007) I tested the effect of agency on passive facilitation and passive harm, and for admiration on active and passive facilitation, contempt on active and passive harm, envy on active harm and passive facilitation and pity on active facilitation and passive harm. Agency was found to be a significant predictor of passive facilitation (Table 16). Contempt was a predictor of active harm (Table 15) and passive harm (Table 17). Envy was found to be a predictor of active harm

(Table 15). Pity was seen to be a predictor of active facilitation (Table 14). Finally, admiration was found to be a predictor of active facilitation (Table 14) and passive facilitation (Table 16).

Table 14

Appearance qualities (that predict masculinity) and emotional reactions that have an effect on active facilitation.

Appearance qualities and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Muscular torso*	0.820	0.212	3.86	0.144	0.0001
Envy**	-0.185	0.080	2.29	-0.089	0.022
Admiration**	0.420	0.045	9.21	0.357	0.0001
Pity**	0.312	0.046	6.68	0.243	0.0001

Note: Regression model: $R^2 = 0.030$, $F(6, 690) = 4.04$, $p < 0.0005^*$ was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.297$, $F(6, 684) = 48.19$, $p < 0.00001^{**}$ was the model run for all mediators on behavioural tendencies.

Table 15

Appearance qualities (that predict masculinity) and emotional reactions that have an effect on active harm.

Appearance qualities and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Chubby torso*	-0.324	0.123	2.63	-0.076	0.009
Contempt**	0.242	0.039	1.20	0.249	0.0001
Envy**	0.307	0.069	4.42	0.215	0.0001

Note: Regression model: $R^2 = 0.011$, $F(6, 689) = 1.70$, $p < 0.1182^*$ was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.150$, $F(6, 684) = 16.05$, $p < 0.00001^{**}$ was the model run for all mediators on behavioural tendencies.

Table 16

Appearance qualities and stereotypical traits (that predict masculinity), and emotional reactions that have an effect on passive facilitation.

Appearance qualities, stereotypical traits and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Muscular torso*	0.629	0.209	3.00	0.119	0.003
Agency**	0.303	0.039	7.78	0.265	0.0001
Admiration**	0.381	0.040	9.35	0.349	0.0001

Note: Regression model: $R^2 = 0.017$, $F(6, 692) = 2.14$, $p < 0.0472^*$ was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.321$, $F(6, 689) = 54.44$, $p < 0.00001^{**}$ was the model run for all mediators on behavioural tendencies.

Table 17

Masculinity and emotional reactions that have an effect on passive harm.

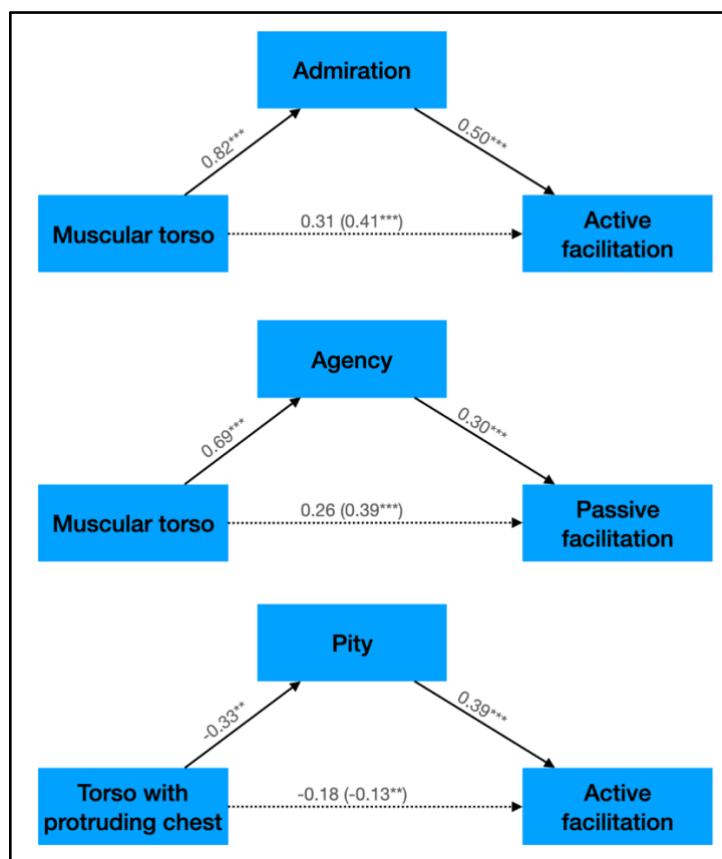
Masculinity and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Masculinity	0.062	0.024	2.57	0.088	0.010
Contempt	0.524	0.040	11.41	0.451	0.0001

Note: Regression model: $R^2 = 0.223$, $F(6, 688) = 33.79$, $p < 0.00001$.

I then ran independent mediations to determine how much of the above effects of appearance qualities on behavioural tendencies were mediated by the mediators of masculinity, agency and emotional reactions (Figure 12). I checked whether the effect of muscular torso was mediated by admiration, and found that the effect of muscular torso on active facilitation vanished as admiration mediated this effect. Furthermore, the mediation of muscular torso on passive facilitation was tested and I found that the direct effect of muscular torso on passive facilitation vanished as the mediation of muscular torso via admiration on passive facilitation mediated the effect. Overall, the only direct effect that remained after the mediations was the effect of chubby torso on active harm.

Figure 12

The effect of appearance qualities that affect masculinity on behavioural tendencies mediated by agency and emotional reactions.



Note: Unstandardized coefficients are reported, with the respective p-values * $p < .05$., ** $p < .01$., *** $p < .001$. Total effects are within the parentheses, and direct effects are adjacent. Full line means direct effect, a dotted line means no direct effect.

In the second model, I tested the effect of appearance qualities (that predict femininity) on behavioural tendencies. In a multiple linear regression analysis I tested the effect of these qualities on behavioural tendencies. I found that a face with make-up significantly negatively affected active facilitation, even if the overall model only approached significance ($R^2 = 0.017$, $F(7, 689) = 1.90$, $p < 0.0671$) (Table 18). Furthermore, pink apparel had a significant negative effect on passive harm (coefficient = -0.764 , $p < 0.005$), where the overall model too was found to be significant ($R^2 = 0.021$, $F(7, 690) = 3.16$, $p < 0.0027$) (Table 19).

I then went on to check whether femininity, communion, and the four emotional reactions were potential predictors of the four behavioural tendencies, and thus, whether these could mediate

the direct effect of appearance qualities on behavioural tendencies. Based on work from Cuddy and colleagues (2007) I tested whether communion would predict active facilitation or active harm. Communion was found to predict active facilitation (Table 18) and active harm (Table 20). In terms of emotions, contempt was found to be a significant predictor of active harm (Table 20) and passive harm (Table 19). Envy was a significant predictor of active harm (Table 20), while pity was found to be a predictor of active facilitation (Table 18). Lastly, admiration was found to be a predictor of active facilitation (Table 18) and passive facilitation (Table 21).

Table 18

Appearance qualities and stereotypical traits (that predict femininity) that have an effect on active facilitation.

Appearance qualities and stereotypical traits	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Face with make-up*	-0.535	0.189	-2.82	-0.110	0.005
Communion**	0.365	0.039	9.29	0.343	0.0001
Pity**	0.185	0.045	4.08	0.145	0.0001
Admiration**	0.375	0.042	8.82	0.319	0.0001

Note: Regression model: $R^2 = 0.017$, $F(7, 689) = 1.90$, $p < 0.0671$ * was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.3594$, $F(6, 689) = 68.58$, $p < 0.00001$ ** was the model run for all mediators on behavioural tendencies.

Table 19

Appearance qualities (that predict femininity) and emotional reactions that have an effect on passive harm.

Appearance qualities and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Pink apparel*	-0.764	0.272	-2.80	-0.077	0.005
Contempt**	0.496	0.047	10.57	0.427	0.0001

Note: Regression model: $R^2 = 0.021$, $F(7, 690) = 3.16$, $p < 0.0027$ * was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.2478$, $F(6, 690) = 41.16$, $p < 0.00001$ ** was the model run for all mediators on behavioural tendencies.

Table 20*Stereotypical traits and emotional reactions that have an effect on active harm.*

Stereotypical traits and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Communion	-0.091	0.029	-3.09	-0.125	0.002
Contempt	0.224	0.039	5.61	0.230	0.0001
Envy	0.295	0.067	4.34	0.206	0.0001

Note: Regression model: $R^2 = 0.1596$, $F(6, 688) = 17.85$, $p < 0.00001$.**Table 21***Emotional reactions that have an effect on passive facilitation.*

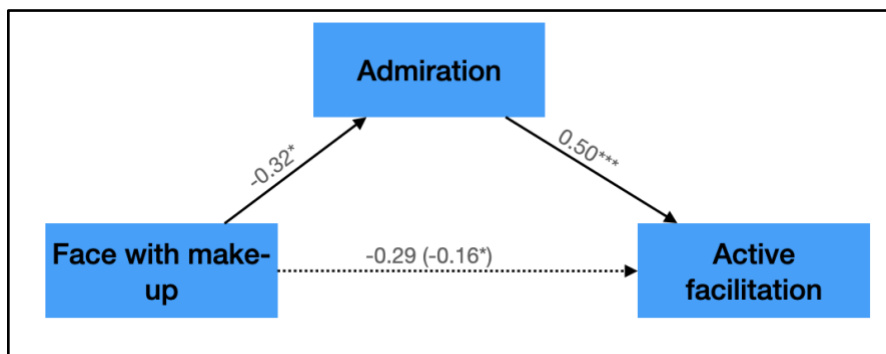
Emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Admiration	0.361	0.040	8.83	0.330	0.0001

Note: Regression model: $R^2 = 0.3467$, $F(6, 691) = 78.55$, $p < 0.00001$.

I then ran independent mediations to determine how much of the above effects of appearance qualities on behavioural tendencies were mediated by the communion and the four emotional reactions (Figure 13). The significant effect of face with make-up on active facilitation vanished, as this effect was mediated by the admiration. The only direct effect that remains is the significantly negative effect of pink apparel on passive harm.

Figure 13

The effect of appearance qualities that affect femininity on behavioural tendencies mediated by emotional reactions.



Note: Unstandardized coefficients are reported, with the respected p-values *p <.05., **p <.01., ***p<.001. Total effects are within the parentheses, and direct effects are adjacent. Full line means direct effect, a dotted line means no direct effect.

In the final model, the effect of appearance qualities (that predict age) on behavioural tendencies were tested. In a multiple linear regression analysis I tested the effect of these qualities on behavioural tendencies. The presence of a nose was a significant negative predictor of active facilitation, and the overall model was significant as well ($R^2 = 0.014$, $F(5, 691) = 2.24$, $p < 0.0491$) (Table 22). Furthermore, white hair was the only significant predictor of active harm (Table 23). Additionally, the presence of a nose and grey hair have a significant direct effect on passive facilitation (Table 24). Finally, I found that the presence of a nose and a short nose have a direct effect on passive harm (Table 25) .

I then went on to check whether age, communion, and the four emotional reactions were potential predictors of the four behavioural tendencies, and found that communion predicted active facilitation (Table 22) and active harm (Table 23). Similarly, contempt was found to be a significant predictor of active harm (Table 23) and passive harm (Table 25). Envy was a significant predictor of active harm (Table 23), while pity was found to be a predictor of active facilitation (Table 22). Lastly, admiration was found to be a predictor of active facilitation (Table 22) and passive facilitation (Table 24).

Table 22

Appearance qualities and stereotypical traits (that predict age), and emotional reactions that have an effect on active facilitation.

Appearance qualities, stereotypical traits and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
The presence of a nose*	-0.420	0.138	-3.05	-0.130	0.002
Communion**	0.360	0.038	9.28	0.339	0.0001
Pity**	0.181	0.045	3.97	0.142	0.0001
Admiration**	0.379	0.042	8.86	0.322	0.0001

Note: Regression model: $R^2 = 0.014$, $F(5, 691) = 2.24$, $p < 0.0491$ * was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.359$, $F(6, 689) = 68.38$, $p < 0.00001$ ** was the model run for all mediators on behavioural tendencies.

Table 23

Appearance qualities and stereotypical traits (that predict age) that have an effect on active harm.

Appearance qualities and stereotypical traits	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
White hair*	0.804	0.325	2.47	0.117	0.014
Communion**	-0.088	0.029	-3.00	-0.121	0.003
Contempt**	0.227	0.040	5.62	0.234	0.0001
Envy**	0.295	0.068	4.32	0.206	0.0001

Note: Regression model: $R^2 = 0.023$, $F(5, 690) = 2.43$, $p < 0.0338$ * was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.159$, $F(6, 688) = 17.38$, $p < 0.00001$ ** was the model run for all mediators on behavioural tendencies.

Table 24

Appearance qualities (that predict age) and emotional reactions that have an effect on passive facilitation.

Appearance qualities and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
The presence of a nose*	-0.564	0.129	-4.35	-0.186	0.0001
Grey hair*	0.822	0.352	2.33	0.070	0.020
Admiration**	0.360	0.170	12.24	0.329	0.0001

Note: Regression model: $R^2 = 0.031$, $F(5, 693) = 5.22$, $p < 0.0001^*$ was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.3458$, $F(6, 691) = 77.55$, $p < 0.00001^{**}$ was the model run for all mediators on behavioural tendencies.

Table 25

Appearance qualities (that predict age) and emotional reactions that have an effect on passive harm.

Appearance qualities and emotional reactions	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
The presence of a nose*	0.452	0.119	3.77	0.169	0.0001
Short nose*	-0.323	0.159	-2.03	-0.089	0.042
Contempt**	0.498	0.047	10.45	0.429	0.0001

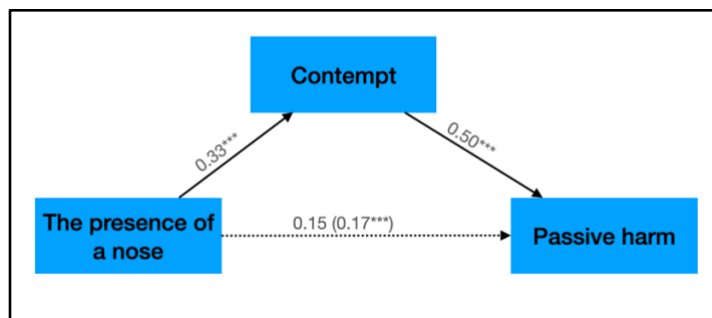
Note: Regression model: $R^2 = 0.027$, $F(5, 692) = 3.39$, $p < 0.0049^*$ was run for all appearance qualities on behavioural tendencies, and $R^2 = 0.2479$, $F(6, 690) = 41.04$, $p < 0.00001^{**}$ was the model run for all mediators on behavioural tendencies.

I then ran independent mediations to determine how much of the above effects of appearance qualities on behavioural tendencies were mediated by the stereotypical trait communion and the four emotional reactions (Figure 14). I found that the effect of presence of a nose on active facilitation remained, and there was no mediation of nose via contempt on active facilitation. Furthermore, it was shown that the effect of the presence of a nose on passive facilitation remained, and that there was a significant mediation of the presence of a nose on passive facilitation via contempt, but that this effect did not mediate the direct effect. Nevertheless,

the direct effect of the presence of a nose on passive harm vanished, as there was a stronger mediation effect of the presence of a nose on passive harm via contempt. Finally, I checked whether the direct effect of grey hair on passive facilitation remained, or whether it was mediated by communion. The direct effect of grey hair on passive facilitation vanished, as this effect was mediated by the effect of grey hair via communion on passive facilitation. Hence, the direct effects that remained were the presence of a nose on active facilitation and passive facilitation, a short nose on passive harm, and white hair on active harm.

Figure 14

The effect of appearance qualities that affect age on behavioural tendencies mediated by emotional reactions.



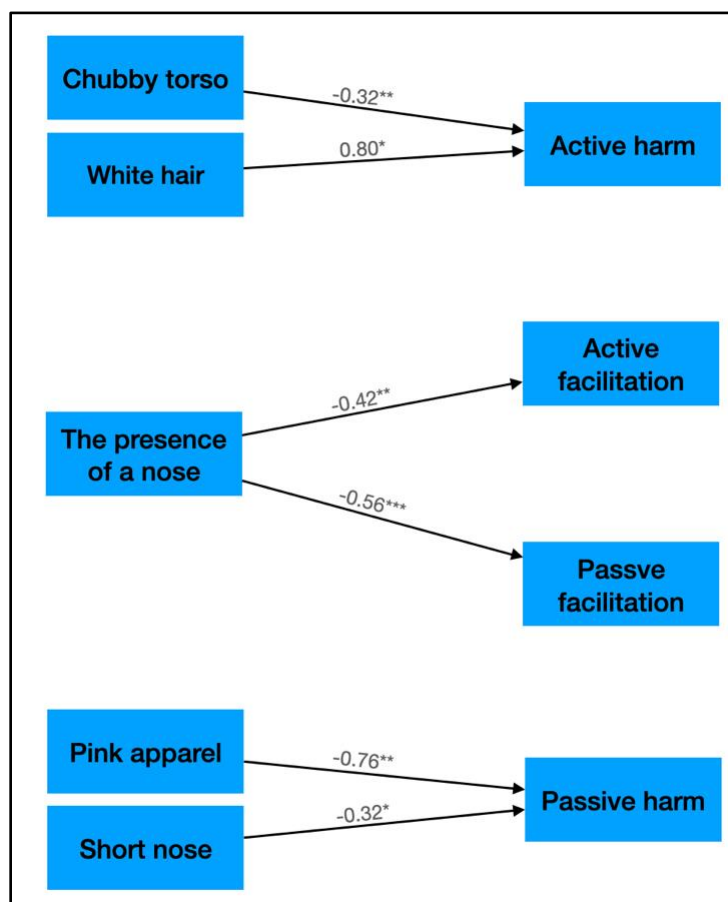
Note: Unstandardized coefficients are reported, with the respected p-values * $p < .05$., ** $p < .01$., *** $p < .001$. Total effects are within the parentheses, and direct effects are adjacent. Full line means direct effect, a dotted line means no direct effect.

In conclusion, for RQ2 I tested which appearance qualities have a direct effect on behavioural tendencies. I found that a chubby torso is a negative predictor of active harm, pink apparel is a negative predictor of passive harm, a short nose is a negative predictor of passive harm, white hair is a positive predictor of active harm, and finally the presence of a nose is a negative predictor of active facilitation and a negative predictor of passive facilitation (Figure 15). Here, none of the hypotheses were confirmed, as the effects were different from what I expected to find (e.g., H4a predicting that a chubby torso would be a positive predictor of active facilitation, when in this study it was a negative predictor of active harm). Similar to observations made for RQ1, I found that the only mediators were stereotypical traits and emotional reactions, but not social categories. The significant mediators were contempt, admiration and agency, whereas communion, pity, envy, masculinity, femininity or age were not found to mediate any effects. However, all

mediations run were significant and mediated the direct effects of the appearance qualities on the behavioural tendencies. For instance, the effects of muscular torso on active facilitation vanished with the effect of muscular torso via admiration. The same results were found for the effects of muscular torso on passive facilitation, which were mediated by agency. While some of the appearance qualities are direct predictors of behavioural tendencies, especially agency and admiration mediated these direct effects.

Figure 15

Appearance qualities that have a significant effect on behavioural tendencies.



Note: Unstandardized coefficients of the regression analyses are reported, with the respected p-values * $p < .05$., ** $p < .01$., *** $p < .001$.

6. Discussion

In this thesis, I studied the effects of robots' appearance qualities on emotional reactions (RQ1) and behavioural tendencies (RQ2), and whether these effects were mediated by the social categories and stereotypical traits people attributed to robots. In this chapter, I will first discuss the findings of the pre-analyses, and then dive into the main findings of this research (i.e., RQ1 and RQ2). I will conclude the chapter with the limitations, recommendations for future research, and the conclusions.

6.1 Interpretation of findings

6.1.1 Discussion pre-analysis: Appearance qualities on social categories

The results of the pre-analyses showed which appearance qualities have an effect on social categories (i.e., masculinity, femininity and age), and on stereotypical traits (i.e., communion and agency). In-line with the hypotheses, it is not just the presence of certain appearance cues (i.e., body manipulators, surface look features and facial features) that predicts masculinity, femininity or age, but more specifically the qualities of these cues that do. For instance, it is not just the presence of a torso that positively predicts masculinity (Perugia et al., 2022), but more specifically the fact that the torso is muscular, chubby, or has visible mechanics (*H1b*). Similarly, it is not just the presence of apparel that positively predicts femininity (Perugia et al., 2022), but specifically whether this apparel is a skirt or a dress (*H1j*) and whether it is pink (*H1j*) and not dark (e.g., blue and dark apparel) (*H1j*). It was also found that a face with make-up is a positive predictor of femininity (*H1h*). Similarly, the presence of eyelashes is a positive predictor of femininity (*H1k*). Finally, in this thesis it is not the presence of hair that negatively predicts age (Kalegina et al., 2018), but rather white or grey hair (*H1p*) that are positive predictors of age. Furthermore, big eyes are a negative predictor of age.

The first interesting result to touch upon is that the regression model of appearance qualities on femininity was the model that explained the most variance ($R^2 = 0.283$, $F(24, 690) = 16$, $p < 0.0001$). This shows that the theory-based appearance qualities selected to predict femininity, are indeed good predictors of femininity. However, interesting discoveries were also made in other regression models, such as the effect of appearance qualities on age. For instance, big eyes are negative predictors of age. These findings are aligned with the “baby schema” (Lorenz, 1943), where big eyes are generally perceived as being more childlike.

It becomes clear that appearance qualities are strong predictors of social categories, sometimes to the extent that the quality of the cue predicts a social category more than the presence of that cue (e.g., the presence of a torso was not a predictor of masculinity, but a *muscular* torso was a positive predictor). The ‘elicited agent knowledge’ theory from Eyssel and colleagues (2012), which originates from the ‘Three-Factor Model of Anthropomorphism’ (Epley et al., 2007), is based on the idea that the extent to which people anthropomorphize objects and nonhuman agents (i.e., ascribe human traits to objects and nonhuman agents) is due to (amongst other psychological factors) the so-called ‘elicited agent knowledge’ (Eyssel et al., 2012; Eyssel & Kuchenbrandt, 2012). The elicited agent knowledge assumes that we are able to anthropomorphize robots since we have larger structures and schemata that we rely on to make sense of unfamiliar objects in our environments. Similarly, I assume that visual cues (e.g., a muscular torso) of robots, activates structures that are associated with related human categories. This could offer an explanation for these observations, and why when participants perceived visual appearance qualities (e.g., muscular torso), they associated the related social category as they would in HHI (e.g., masculinity).

Another explanation could have been that participants were not able to see the specific qualities in the images that they viewed of the robots, meaning that they only spotted the qualities that stood out to them the most (e.g., pink apparel, which they then associated with femininity), which will further be discussed in the limitations. The following table gives an overview of which hypotheses were rejected and which ones were partially accepted.

Table 26

Hypotheses for pre-analysis: appearance qualities on social categories.

Hypotheses	Accepted (A) / partially accepted (PA) / rejected (R)	Specific qualities direction of effects (if (partially) accepted)
<i>H1a</i> : The presence of short hair positively predicts masculinity.	R	
<i>H1b</i> : The presence of a torso (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, curvy, or with visible mechanics) predicts masculinity.	PA	Muscular, chubby and torso with visible mechanics are positive predictors of

<p><i>H1c</i>: The presence of arms (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics) predicts masculinity.</p> <p><i>H1d</i>: The presence of hands (e.g., moveable, or able to grasp objects) predicts masculinity.</p> <p><i>H1e</i>: The presence of fingers (e.g., long, short, or chubby) predicts masculinity.</p> <p><i>H1f</i>: The presence of legs (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics) predicts masculinity.</p>	<p>masculinity. Slender torso is a negative predictor of masculinity. None of the other qualities (e.g., long, short etc.) are predictors of masculinity.</p>
<p><i>H1c</i>: The presence of arms (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics) predicts masculinity.</p>	<p>R</p>
<p><i>H1d</i>: The presence of hands (e.g., moveable, or able to grasp objects) predicts masculinity.</p>	<p>R</p>
<p><i>H1e</i>: The presence of fingers (e.g., long, short, or chubby) predicts masculinity.</p>	<p>R</p>
<p><i>H1f</i>: The presence of legs (e.g., long, short, metallic, realistic (i.e., life-like, human-like), muscular, slender, chubby, or with visible mechanics) predicts masculinity.</p>	<p>PA</p> <p>The presence of legs is a positive predictor of masculinity. None of the qualities (e.g., long, short, metallic etc.) are predictors of masculinity.</p>
<p><i>H1g</i>: The presence of skin (e.g., metallic, realistic (i.e., life-like, human-like), black, brown, white, pink, yellow, or another colour) predicts femininity.</p>	<p>R</p>
<p><i>H1h</i>: The presence of a face with make-up (e.g., red lips) positively predicts femininity.</p>	<p>PA</p> <p>Face with make-up is a positive predictor of femininity. Red lips is not a predictor of femininity.</p>
<p><i>H1i</i>: The presence of full lips positively predicts femininity.</p>	<p>R</p>
<p><i>H1j</i>: The presence of apparel (e.g., t-shirt, skirt, trousers, dress, jacket, suit, mask, helmet, armour, tight, loose, formal, informal, pink, blue, dark, white, or another colour) predicts femininity.</p>	<p>PA</p> <p>Skirt, dress and pink apparel are positive predictors of femininity. Blue and dark are negative predictors</p>

<p><i>Hlk</i>: The presence of eyelashes (e.g., long, short) predicts femininity.</p> <p><i>Hll</i>: The presence of head hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like) blond/ yellow, dark, grey, white, or another colour) predicts femininity.</p> <p><i>Hlm</i>: The presence of eyebrows (e.g., thick, thin, metallic, realistic (i.e., life-like, human-like), straight, curved, raised, or frowning) predicts femininity.</p>	<p>of femininity. None of the other qualities (e.g., jacket, helmet etc.) are predictors of femininity.</p> <p>PA</p> <p>The presence of eyelashes is a positive predictor of femininity. None of the qualities (i.e., long, short) are predictors of femininity.</p> <p>R</p> <p>R</p>
<p><i>Hln</i>: The presence of eyes (e.g., big, small, rounded, elongated, expressive, moveable, metallic, realistic (i.e., life-like, human-like), mechanical, or with big pupils) predicts age.</p> <p><i>Hlo</i>: The presence of a nose (e.g., long, short, rounded, pointed, narrow, wide) predicts age.</p> <p><i>Hlp</i>: The presence of hair (e.g., long, short, metallic, realistic (i.e., life-like, human-like), blond/ yellow, dark, grey, white, or another colour) predicts age.</p>	<p>PA</p> <p>Big eyes is a negative predictor of age. None of the other qualities (e.g., moveable, metallic) are predictors of age.</p> <p>PA</p> <p>The presence of a nose is a positive predictor of age. A short nose is a negative predictor of age. None of the other qualities (e.g., narrow, wide) are predictors of age.</p> <p>PA</p> <p>White and grey hair are positive predictors of age. None of the other qualities (e.g., metallic, realistic) are predictors of age.</p>

H1q: The presence of a **head** (e.g., big, small, elongated, short, rounded, squared, spherical, oval, realistic (i.e., life-like, human-like), mechanical, a screen/ tablet, or a camera) predicts age. R

H1r: The presence of a **face** (e.g., realistic (i.e., life-like, human-like), metallic, mechanical, chubby, sunken, expressive, wrinkled, bearded / with moustache, or with make-up) predicts age. R

H1s: The presence of a **mouth** (e.g., big, small, squared, smiling, minimalistic, moveable, red, featuring full lips, or featuring thin lips) predicts age. R

Note: Yellow shows which hypotheses were partially accepted, red shows which hypotheses were rejected. Green highlights the hypotheses that predict masculinity, blue highlights the hypotheses that predict femininity, and orange highlights the hypotheses that predict age.

6.1.2 Discussion pre-analysis: Appearance qualities and social categories on stereotypical traits

In the second pre-analysis I examined which appearance qualities predicted the stereotypical traits of communion and agency, and whether the social categories attributed to the robot (i.e., femininity, masculinity, and age) have a mediating role in this relationship. Similar to Perugia and colleagues (2023), I expected to find direct effects specifically of the appearance *qualities* (i.e., for masculinity these are the presence of a torso that is muscular, chubby, slender, or has visible mechanics and the presence of legs, for femininity these are face with make-up, skirt, dress, blue, dark and pink apparel, and the presence of eyelashes, and for age these are big eyes, the presence of a nose, white and grey hair (Table 27)) to have a direct effect on stereotypical traits (i.e., appearance qualities that predict masculinity have a direct positive effect on agency, appearance qualities that predict femininity and age have a direct positive effect on communion) as well as indirect effects on stereotypical traits mediated by social categories (i.e., appearance qualities that predict masculinity have an indirect positive effect on agency mediated by masculinity, appearance qualities that predict femininity and age have an indirect positive effect on communion mediated by femininity and age respectively).

Here, amongst other effects (Table 27), I found that a muscular torso has a positive direct effect on agency (*H2a*), chubby torso and torso with visible mechanics had a negative effect on agency, and pink apparel and grey hair have a positive direct effect on communion (*H2b* and *H2c*).

These findings are in-line with my predictions, which is that the appearance qualities that predict a social category (e.g., pink apparel on femininity), will have the same effect on the stereotypical traits as the social category would (e.g., pink apparel positively predicts communion). The reason for these effects could be that the appearance qualities inherently carry essential information about the social categories they depict. Consequently, these appearance qualities directly trigger stereotypical traits that are commonly associated with those social categories.

Two main additional findings were that pink apparel has a positive direct effect on communion, and grey hair has a positive direct effect on communion. These results might be interpreted with humanlikeness, as Perugia and colleagues (2022) discovered that humanlikeness is an important factor when it comes to social categorisation. In this thesis this could mean that people ascribe gender and age to robots based on which social category the appearance quality reminds them of, and more importantly, that the appearance quality holds intrinsic information of the social category, such that it is a direct reminder not only of social category (e.g., femininity) but rather of the associated stereotypical trait (e.g., communion). Similarly, grey hair seems to be very reminiscent of older age and more importantly, communion.

On top of these direct effects, I also found two indirect effects: a positive indirect effect of muscular torso on agency via masculinity (*H2a*), and a positive indirect effect of pink apparel on communion via femininity (*H2b*). These two indirect effects show that a muscular torso of a robot evokes the perception that the robot is agentic, as muscular torso is linked to masculinity which is linked to agency. Pink apparel, on the other hand, elicits the feeling that the robot has more communion, as pink apparel is linked to femininity which relates to communion. These effects are in-line with previous literature that suggest that agency is related to masculinity whereas communion is associated with femininity (Bernotat et al., 2021; Eyssel & Hegel, 2012). It becomes clear here that certain physical attributes (e.g., muscular torso) and clothing items (e.g., pink apparel) that are culturally and socially associated with masculinity and femininity, go on to influence similar stereotypical traits (agency and communion) evoked by the social categories, as the social categories too are culturally and socially associated with certain stereotypes. For qualities that predicted age, no significant indirect effects were found, which is why only partial proof of *H2c* was found (i.e., grey hair has a direct positive effect on communion).

Another important observation made in these pre-analyses was that there are direct and

indirect effects of appearance qualities on stereotypical traits, but that not all appearance qualities had both direct and indirect effects. The results of this thesis indicate that both direct and indirect effects arise when the appearance quality is strongly associated with the social category. For instance, muscular torso had both a positive direct and indirect effect on agency, compared to slender torso which had neither direct nor indirect effects. Reasons for this could be that muscular torso is stronger associated with masculinity than slender torso is, especially as slender torso was found to be a negative predictor of masculinity in the first pre-analysis. Similarly, the presence of legs was not found to have a direct effect nor indirect effect on agency, which is likely due to the fact that it is not strongly enough associated with masculinity and therefore does not lead to the perception of agency. Moreover, a reason why the presence of legs did not have a direct or indirect effect on agency might be because this is an appearance quality that is associated with more social categories (e.g., femininity). In comparison, torso with visible mechanics had direct and indirect effects on agency. The effects of torso with visible mechanics on agency might, however, be negative due to the fact that visible mechanics might reduce the humanlikeness that people feel towards this appearance quality, which in turn reduces the perceived agency of the robot. Another reason for this negative effect of visible mechanics on agency might be that visible mechanics lead to less freedom to perform various tasks, and therefore might be associated with less agentic robots. As agency refers to a robot's perceived ability to perform tasks and achieve goals (Perugia et al., 2023), visible mechanics and wires might not be associated with an agency per-se.

Similarly, a chubby torso was also found to be a negative direct predictor of agency. However, this observation points to a different fact, which might be that chubby torso is a potential predictor of age rather than of masculinity. While a chubby torso is a predictor of masculinity, in this thesis, no exploratory analyses were run regarding the effects that chubby torso could have on age. While it might be normal that the appearance quality alone means something different than in combination with the social category, in retrospect, chubby might be more associated with childlikeness. Going back to the “baby schema”, some of the features of childlikeness include overall rounder bodies (Lorenz, 1943). This would also explain why a chubby torso is a negative predictor of agency, as child-like robots are seen as being more helpless and vulnerable than adult robots (Dunstan & Hoffman, 2023). This shows how much information the appearance qualities hold on their own. The following table gives an overview of which hypotheses accepted.

Table 27

Hypotheses for pre-analysis: appearance qualities and social categories on stereotypical traits (communion and agency).

Hypotheses	Accepted (A) / partially accepted (PA) / rejected (R)	Specific qualities direction of effects (if (partially) accepted)
<p><i>H2a:</i> Appearance qualities that predict masculinity (i.e., the presence of a torso that is muscular, chubby, slender, or has visible mechanics and the presence of legs) have a positive direct effect on agency and a positive indirect effect on agency mediated by masculinity.</p>	PA	<p>Muscular torso has a positive direct effect on agency, and a positive indirect effect on agency via masculinity.</p> <p>Chubby torso has a negative direct effect on agency, and a positive indirect effect on agency via masculinity.</p> <p>Torso with visible mechanics had a negative direct effect on agency, and a positive indirect effect on agency via masculinity.</p>
<p><i>H2b:</i> Appearance qualities that predict femininity (i.e. face with make-up, skirt, dress, blue, dark and pink apparel, and the presence of eyelashes) have a positive direct effect on communion and a positive indirect effect on communion mediated by femininity.</p>	PA	<p>Pink apparel has a positive direct effect on communion, and a positive indirect effect on communion via femininity.</p>
<p><i>H2c:</i> Appearance qualities that predict age (i.e., big eyes, the presence of a nose, white and grey hair) have a positive direct effect on communion and a positive indirect effect on communion mediated by age.</p>	PA	<p>Grey hair has a positive direct effect on communion.</p>

Note: Yellow shows which hypotheses were partially accepted. Green highlights the hypotheses that predict masculinity, blue highlights the hypotheses that predict femininity, and orange highlights the hypotheses that predict age.

6.1.3 Discussion RQ1: Appearance qualities on emotional reactions

In the first main analysis, the effects of appearance qualities on emotional reactions were examined. A muscular torso was found to be a positive direct predictor of admiration (*H3a*), blue and dark apparel were negative direct predictors of envy, face with make-up was a negative direct predictor of admiration, big eyes was a positive direct predictor of contempt and pity (*H3c*), and the presence of a nose was a positive direct predictor of contempt (Figure 11 in the results section). Less hypotheses were confirmed in these analyses, as the significant direct effects of appearance qualities on emotional reactions differed from expectations, and the same held true for the hypothesised mediations of social categories and stereotypical traits (see Table 28 for a full overview). For the appearance qualities where the effects were mediated, these were only mediated by age and agency (e.g., positive effect of muscular torso on agency and through agency on admiration, and positive effect of white hair on age and via age on contempt), and not communion, femininity nor masculinity.

The fact that femininity and masculinity did not mediate any effects seem to stem from appearance qualities being gendered in the first place, such that they convey which social category they represent, and thereby mitigating any effects of social categories. While femininity was not found to predict any emotional reactions, the qualities that predicted femininity in the first pre-analysis (e.g., face with make-up (+), blue and dark apparel (-)) were significant predictors of emotional reactions (e.g., face with make-up was a negative direct predictor of admiration, and blue and dark apparel were negative direct predictors of envy). The same was found for the appearance qualities that predict masculinity: these qualities had a direct strong effect on emotional reactions (e.g., muscular torso is a positive direct predictor of admiration) but this effect was not mediated by masculinity. The reason why muscular torso had a positive effect on admiration whereas face with make-up did might relate to findings from Cuddy and colleagues (2007) and Fiske and colleagues (2002). They discovered that usually the social groups that elicit admiration are the in-groups (i.e., the societal norm), so those that are high in both warmth and competence. Here, it is interesting to notice that while a muscular torso, which was seen to predict masculinity in the first pre-analysis, positively predicts admiration, face with make-up, which was seen to predict femininity in the first pre-analysis, negatively predicts admiration. I would have expected to find that face with make-up too elicits admiration. As admiration is usually elicited by in-groups, the positive relationship between muscular torso and admiration and the negative relationship

between face with make-up and admiration seem to suggest that male robots, especially those that fit the aesthetic stereotype of masculinity, can be considered in-groups, whereas female robots, especially those that fit the aesthetic stereotype of femininity, can be considered out-groups.

Another interesting finding was that while femininity and masculinity were not mediators of any effects, age was. The positive direct effect of the white hair on contempt vanished as this effect was mediated by the positive effect of white hair via age on contempt. The direct positive effect of big eyes on contempt was partially mediated by age, which lead to a negative indirect effect of big eyes on contempt via age. However, these mediations were not full mediations as the direct effect of big eyes on contempt remained. This shows how strong the direct effect of big eyes on emotional reactions (i.e., pity and contempt) is. Especially for pity, there was not even a partial mediation by age, and big eyes themselves evoked this emotional reaction. A reason for this could be that big eyes, and therefore younger age, might evoke caregiving responsibilities in a person, which can then lead to the person taking on the role of protector and caretaker (CaNAmero et al., 2006). These perceived responsibilities can lead to the person's nurturing instincts being activated, and for them to feel certain emotional reactions (e.g., pity) towards the robots to help it (Kirkpatrick, 1992; Rabb et al., 2022). Big eyes seem to be a strong indicator of childlikeness as they can carry the meaning about age on their own without the need for a mediation by age.

Another unexpected finding in this thesis was that communion was not a mediator of any of the direct effects, however, agency was. For instance, torso with visible mechanics had a negative indirect effect on envy via agency. This could mean that agency, and therefore the ability of the robot to act, is more crucial in mediating emotional reactions than communion is. Relating these findings back to the SCM of Fiske and colleagues (2002), agency might take precedence over communion when it comes to ensuring our personal safety, which might explain why it relates more closely to emotional reactions, when robots are evaluated by laypeople. It was found that muscular torso had a positive direct effect on admiration, and a positive indirect effect on admiration via agency.

As I expected the effect of the appearance qualities would be mediated by social categories and stereotypes together, none of the hypotheses were fully accepted. Overall, looking at the magnitude of the direct effects that were not mediated, the strongest predictors of emotional reactions of the appearance qualities that predict masculinity were muscular torso (which had a

positive effect on admiration), for femininity face with make-up (which had a negative effect on admiration), and for age big eyes (which had positive effects on contempt and pity). Comparing the direct effects to the indirect effects, for instance to the positive effect of muscular torso on admiration via agency, it becomes clear that these direct effects had a greater magnitude than the mediating effects. The same was found for the indirect effect of big eyes on contempt via age, where the direct effect did not vanish. This shows how strong the direct effects of appearance qualities on emotional reactions actually are, and while age and agency mediate some of the direct effects of the appearance qualities, appearance qualities themselves carry enough information (e.g., regarding the social categories) to elicit emotional reactions. For instance, the effect of muscular torso remained significant on admiration despite any mediation by agency, which proves that a muscular torso itself (which is a strong predictor of masculinity) is closely associated to admiration. The following table gives an overview of which hypotheses were rejected and which ones were partially accepted.

Table 28

Hypotheses for RQ1: Appearance qualities on emotional reactions.

Hypotheses	Accepted (A) / partially accepted (PA) / rejected (R)	Specific qualities direction of effects
<p><i>H3a:</i> Appearance qualities that predict masculinity (i.e., the presence of a torso that is muscular, chubby, slender, or has visible mechanics and the presence of legs) have a positive direct effect on envy and admiration and a positive indirect effect on envy and admiration mediated by agency and masculinity.</p>	PA	<p>Muscular torso has a positive direct effect on admiration, and a positive indirect effect on admiration via agency.</p> <p>Torso with visible mechanics has a negative indirect effect on envy via agency.</p>

<p><i>H3b</i>: Appearance qualities that predict femininity (i.e. face with make-up, skirt, dress, blue, dark and pink apparel, and the presence of eyelashes) have a positive direct effect on pity and admiration and a positive indirect effect on pity and admiration mediated by communion and femininity.</p>	R	<p>Blue and dark apparel have negative direct effects on envy.</p> <p>Face with make-up has a negative direct effect on admiration.</p>
<p><i>H3c</i>: Appearance qualities that predict age (i.e., big eyes, the presence of a nose, white and grey hair) have a positive direct effect on pity and admiration and a positive indirect effect on pity and admiration mediated by communion and age.</p>	PA	<p>Big eyes have a positive direct effect on pity.</p> <p>Big eyes have a positive direct effect on contempt, and a negative indirect effect on contempt via age.</p> <p>The presence of a nose has a positive direct effect on contempt, and a positive indirect effect on contempt via age.</p> <p>White hair has a positive indirect effect on contempt via age.</p>

Note: Yellow shows which hypotheses were partially accepted, and red shows which hypotheses were rejected. Green highlights the hypotheses that predict masculinity, blue highlights the hypotheses that predict femininity, and orange highlights the hypotheses that predict age.

6.1.4 Discussion RQ2: Appearance qualities on behavioural tendencies

In RQ2, I tested which appearance qualities have a direct effect on behavioural tendencies. A chubby torso was found to be a negative predictor of active harm, a short nose a negative predictor of passive harm, white hair a positive predictor of active harm, the presence of a nose a negative predictor of active facilitation and a negative predictor of passive facilitation, and finally pink apparel a negative predictor of passive harm (Figure 15 in the results section).

An unexpected finding here was that the appearance qualities that predict age were the ones to have the most effect on behavioural tendencies (i.e., all but chubby torso and pink apparel were effects coming from age related appearance qualities). Another surprising finding here was that the appearance qualities that indicate older age induce negative behavioural tendencies (e.g., white hair having a positive direct effect on active harm), whereas the indicators of younger age (e.g., a shorter nose being a negative direct predictor of passive harm) induces positive behavioural tendencies (i.e., overall lower passive and active harm). Similar to what was mentioned earlier, chubby torso seems to be a predictor of age and not of masculinity. This would explain why a chubby torso is a negative predictor of active harm, as the caregiver instinct of the person would want to protect rather than harm a child-like robot. In this thesis, a chubby torso was rated highest (between 0 - 1) for the robots Adata (*mean* = 0.8), Discorobo (*mean* = 0.6), and Pino (*mean* = 0.5). Adata was on average perceived as being 12.3 years old, Discorobo was on average perceived as being 8.4 years old, and Pino was on average perceived as being 36.6 years old.

Furthermore, in these analyses none of the hypotheses were confirmed, the main reason being that the direction of the effects of appearance qualities on behavioural tendencies as well as the mediations I expected to see were different from what was found in this thesis (see Table 29 for a detailed overview). Another reason why the original hypotheses were not confirmed is that, similar to RQ1, there were no complete mediations, but the mediators were either stereotypical traits or emotional reactions but never social categories. The mediators were contempt, admiration and agency, however not communion, pity, envy, masculinity, femininity nor age. The reason why these mediators were found, especially admiration and agency, might relate back to findings from the SCM (Fiske et al., 2002) and theories of intergroup relations (Glick & Fiske, 1999). Agency might be a mediator for the same reason as was discussed before, that agency might take precedence over communion when it comes to ensuring our personal safety (Fiske et al., 2002).

Similarly, admiration can serve as a means to highlight the social groups that possess certain advantages in society, such as greater economic safety. These advantages may be admired by both in-groups (e.g., men) and out-groups (e.g., women), such as men's elevated status within society (Glick & Fiske, 1999). These advantages are not solely sought after by the in-group (e.g., men), but rather they are relied upon by out-group members (e.g., women). Traditionally, this reliance has often been associated with economic security that women have frequently had to depend on men for (Glick & Fiske, 1999). Therefore, admiration can reveal where the greatest sense of security lies, which underscores its importance.

Presumably the reason why social categories did not mediate any effects is that the appearance qualities themselves gave enough information regarding which social categories they predict. For the mediations that were run, all were found to be significant. For instance, the positive effect of muscular torso on passive facilitation was fully mediated by agency. Here, it became clear that indirect effects had a greater magnitude than the direct effects did on behavioural tendencies, which shows the role that the admiration has. The same results were found for the positive effect of muscular torso on active facilitation, where the effect vanished as the positive effect of muscular torso via admiration on active facilitation was stronger. This last result is interesting to connect with what was found for face with make-up, which is that there is a negative effect of face with make-up on admiration, and through admiration a negative effect on active facilitation. While muscular torso is positively associated with admiration and, through it, with active facilitation, face with make-up is negatively associated with admiration and, through it, with active facilitation. As admiration is usually elicited by in-groups, these findings again seem to show that male robots, those that fit the aesthetic stereotype of masculinity, can be considered in-groups, whereas female robots, those that fit the aesthetic stereotype of femininity, can be considered out-groups. Even more so, it seems that fitting this aesthetic stereotype determines whether or not you are going to be facilitated in an interaction, as to say whether you are going to be helped and protected. Furthermore, it becomes clear that admiration, in-line with the BIAS map (Cuddy et al., 2007), predicts active facilitation.

While some of the appearance qualities were direct predictors of behavioural tendencies (e.g., short nose is a negative direct predictor of passive harm), especially admiration was found to be an important mediator of many effects. Relating these findings back to literature, we notice that the BIAS map found the concept of “emotional priority”, which means that emotional

reactions have a greater effect on behavioural tendencies than stereotypical traits have on behavioural tendencies (Cuddy et al., 2007; Mieczkowski et al., 2019). Here, we might therefore actually say that emotions have a greater effect on behaviours than social categories and stereotypical traits have. Another explanation could be that participants evaluating the robots in this thesis were most likely primed by what they know about them in the media. Therefore, the robots might have evoked emotional reactions based on prior subconscious perceptions, and regardless of how participants rated social categories or stereotypes of robots, emotional reactions are unfiltered responses that represent how people actually feel towards robots. The following table gives an overview of which hypotheses were rejected and why.

Table 29

Hypotheses for RQ2: Appearance qualities on behavioural tendencies.

Hypotheses	Accepted (A) / partially accepted (PA) / rejected (R)	Specific qualities direction of effects (if rejected)
<p><i>H4a:</i> Appearance qualities that predict masculinity (i.e., the presence of a torso that is muscular, chubby, slender, or has visible mechanics and the presence of legs) have a positive direct effect on active and passive facilitation, and active harm and a positive indirect effect on active and passive facilitation, and active harm mediated by admiration, envy, agency, and masculinity.</p>	R	<p>Muscular torso has a positive indirect effect on active facilitation via admiration.</p> <p>Muscular torso has a positive indirect effect on passive facilitation via agency.</p> <p>Chubby torso has a negative direct effect on active harm.</p>
<p><i>H4b:</i> Appearance qualities that predict femininity (i.e. face with make-up, skirt, dress, blue, dark and pink apparel, and the presence of eyelashes) have a positive direct effect on active facilitation, passive harm and passive facilitation and a positive indirect effect on active facilitation, passive harm and passive facilitation mediated by pity, admiration, communion, and femininity.</p>	R	<p>Face with make-up has a negative indirect effect on active facilitation via admiration.</p> <p>Pink apparel has a negative direct effect on passive</p>

	harm.
<p><i>H4c</i>: Appearance qualities that predict age (i.e., big eyes, the presence of a nose, white and grey hair) have a positive direct effect on active facilitation, passive harm and passive facilitation and a positive indirect effect on active facilitation, passive harm and passive facilitation mediated by pity, admiration, communion and femininity.</p>	<p>R</p> <p>A short nose has a negative direct effect on passive harm.</p> <p>White hair has a positive direct effect on active harm.</p> <p>The presence of a nose has a negative direct effect on active facilitation and passive facilitation, and a positive indirect effect on passive harm via contempt.</p>

Note: Red shows which hypotheses were rejected. Green highlights the hypotheses that predict masculinity, blue highlights the hypotheses that predict femininity, and orange highlights the hypotheses that predict age.

6.1.5 General discussion

In the pre-analyses and analyses of RQ1 and RQ2, interesting observations were made. Firstly, I found that appearance qualities, which socio-culturally associate with social categories, strongly predicted social categories in this thesis as well (e.g., masculinity was associated with a muscular torso, femininity with pink apparel, and age with grey hair). Furthermore, I discovered an unexpected finding regarding the mediation effect of social categories. While social categories partially mediated the effects of appearance qualities on stereotypical traits, further analysis of RQ1 revealed that gender no longer mediated any effects, leaving only age as a mediator. Similarly, agency, not communion, was the sole stereotypical trait mediating any direct effects. In RQ2, it became evident that agency and admiration were the primary mediators, while all other social categories (including gender and age) and stereotypical traits became irrelevant.

Seeing that social categories do not mediate any effects on behavioural tendencies it becomes clear that appearance qualities themselves carry a lot of meaning when it comes to robot perception, not just in terms of gender and age attribution, but also in terms of the elicitation of stereotypical traits, emotional reactions, and behavioural tendencies. Understanding which appearance qualities have an effect on emotions and behaviours is especially important if we want to understand and investigate thoroughly whether the way we design robots and imbue certain stereotypical aesthetic features into them could lead to a different perception and treatment. Using Sparrow's (2021) argumentation, knowing that a robot with white or grey hair is more likely to be actively harmed, or that robots with make-up lead to less admiration, tells us something about our human nature, and how the biases and negative tendencies we censor when interacting with other humans (e.g., older people or women) might resurface when we interact with robots (e.g., robots perceived as older or more feminine).

6.2 Limitations

While important results were found, there are some limitations that should be acknowledged. First, in the pre-analysis, I selected the appearance qualities manually (e.g., long, short, metallic, human-like hair) by going through the ABOT dataset and taking notes of which qualities I saw. Therefore, certain qualities might have been excluded in this research that could have been important predictors of social categories, and later on of stereotypical traits, emotional reactions, and

behavioural tendencies. Additionally, the order in which the questions regarding the appearance qualities were asked (e.g., asking participants to rate the robot head to toe) could have resulted in an order bias, as the presentation order of the questions was not randomised. Along with this, as participants viewed images of robots taken from a certain perspective and with different resolutions, certain appearance qualities might have not been visible enough (e.g., too small/blurry).

Furthermore, all analyses that were run in this thesis were based on theory, meaning that only the effects that were highlighted by previous research were examined. Hence, for appearance qualities that predicted masculinity, agency was used as a mediator, and vice versa the mediator communion was used for qualities that predict femininity and age. However, had I run more exploratory analyses to determine whether communion could mediate the effects of qualities that predict masculinity, and agency mediate the effects of qualities that predict femininity and age, different conclusions might have been made. In the same line of thought, it would have been important to control for femininity and vice versa for masculinity when running the analyses on the research questions. For instance, when wanting to understand whether femininity predicted emotional reactions, I should have probably also controlled for the potential effects of masculinity. Furthermore, it would have been important to test which appearance qualities have an effect on different social categories, not to test those from previous literature. This is especially important as it became clear that, for instance, chubby torso might be a predictor of age rather than masculinity.

Moreover, had there been time, it would have been extremely interesting to study the effects of participants' individual characteristics (e.g., tendency to anthropomorphise, age, gender, benevolent and hostile sexism) on the dependent variables in the study (e.g., robot's perceived age and gender, stereotypical traits, emotional reactions, and behavioural tendencies). This could have cast a light on whether some of the sexist and ageist results we discovered (e.g., face with make-up negatively predicting admiration and, via admiration, active facilitation) could be related to participants' identity and belief system (e.g., benevolent and hostile sexism).

6.3 Conclusion

The present research showed that it is not so much the appearance cues (e.g., the presence of a torso) that lead to social categorisation, but rather appearance *qualities* (e.g., *muscular* torso) that predict social categories (e.g., masculinity) and stereotypical traits (e.g., agency). Appearance qualities were found to go as far as to have an effect on emotional reactions and behavioural tendencies (e.g., white hair positively predicting active harm). Both theory-based results (e.g., muscular torso having a positive direct effect on admiration) as well as new, unexpected results (e.g., face with make-up having a negative direct effect on admiration) were found. What became apparent in this research is that stereotypical traits (in particular agency) mediated some of the direct effects of appearance qualities on emotional reactions and behavioural tendencies, whereas gender did not. This highlights one of the main findings in this thesis, which is that appearance qualities hold a magnitude of information regarding the social category that they represent, which can directly shape how we perceive and behave towards robots.

7. Acknowledgements

I would like to express my sincere gratitude to Giulia Perugia, my first supervisor from TU/e, for her unwavering support and invaluable academic insights while writing my thesis. Her guidance has been tremendously valuable and greatly appreciated. Additionally, I would like to express my heartfelt appreciation to Stefano Guidi, my secondary supervisor from the University of Siena, for his valuable feedback, insightful contributions, and overall enthusiasm in the completion of this thesis.

Additionally, I want to thank my friends and family (specifically to Ines, my mom, dad and sisters) for being incredibly patient with me and being my number one fans. Your knowledge and guidance have been priceless. Last but not least, I want to thank Nes for being my greatest supporter throughout this journey and for always having my back, I could not have done this without you!

8. References

- Abdi, J., Al-Hindawi, A., Ng, T., & Vizcaychipi, M. P. (2018). Scoping review on the use of socially assistive robot technology in elderly care. *BMJ Open*, 8(2).
- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of personality and social psychology*, 93(5), 751.
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, 6(2), 269–278.
- Bernotat, J., Eyssel, F., & Sachse, J. (2021). The (fe) male robot: how robot body shape impacts first impressions and trust towards robots. *International Journal of Social Robotics*, 13, 477-489.
- Biddle, B.J. (1979). Role theory: Expectations, identities, and behaviours. *New York: Academic Press*.
- Breazeal C., & Foerst, A. (1999). Schmoozing with robots: exploring the boundary of the original wireless network. In: *Proceedings of the 3rd international cognitive technology conference (CT-99)*, pp 375–390. San Francisco, CA.
- CaNAmero, L., Blanchard, A. J., & Nadel, J. (2006). Attachment bonds for human-like robots. *International Journal of Humanoid Robotics*, 3(03), 301-320.
- Choi, N., Fuqua, D. R., & Newman, J. L. (2009). Exploratory and confirmatory studies of the structure of the Bem Sex Role Inventory short form with two divergent samples. *Educational and Psychological Measurement*, 69(4), 696-705.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of personality and social psychology*, 92(4), 631.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology*, 40, 61-149.
- Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *Journal of personality and social psychology*, 50(5), 925.
- Dahlgren, K. (1985). The cognitive structure of social categories. *Cognitive science*, 9(3), 379-398.
- Dunstan, B. J., & Hoffman, G. (2023). Social Robot Morphology: Cultural Histories of Robot Design. In *Cultural Robotics: Social Robots and Their Emergent Cultural Ecologies* (pp. 13-34). Cham: Springer International Publishing.

- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, 114(4), 864.
- Eyssel, F., & Hegel, F. (2012). (s)he's got the look: Gender stereotyping of robots 1. *Journal of Applied Social Psychology*, 42(9), 2213-2230.
- Eyssel, F., Kuchenbrandt, D., Hegel, F., & De Ruiter, L. (2012, September). Activating elicited agent knowledge: How robot and user features shape the perception of social robots. In *2012 IEEE RO-MAN: The 21st IEEE international symposium on robot and human interactive communication* (pp. 851-857). IEEE.
- 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.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology*, 82(6), 878.
- George, P. A., & Hole, G. J. (2000). The role of spatial and surface cues in the age-processing of unfamiliar faces. *Visual Cognition*, 7(4), 485-509.
- 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, 73.
- Glick, P., & Fiske, S. T. (1999). The ambivalence toward men inventory: Differentiating hostile and benevolent beliefs about men. *Psychology of women quarterly*, 23(3), 519-536.
- Guidi, S., Boor, L., van der Bij, L., Foppen, R., Rikmenspoel, O., & Perugia, G. (2022, December). Ambivalent Stereotypes Towards Gendered Robots: The (Im)mutability of Bias Towards Female and Neutral Robots. In *International Conference on Social Robotics* (pp. 615-626). Cham: Springer Nature Switzerland.
- Helman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. (2014). Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychological and Personality Science*, 5(7), 815-824.
- Hummert, M. L., Garstka, T. A., & Shaner, J. L. (1997). Stereotyping of older adults: The role of target facial cues and perceiver characteristics. *Psychology and aging*, 12(1), 107.
- Kalegina, A., Schroeder, G., Allchin, A., Berlin, K., & Cakmak, M. (2018, February). Characterizing the design space of rendered robot faces. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 96-104).
- Kaufmann, M. C., Krings, F., & Szesny, S. (2016). Looking too old? How an older age appearance reduces chances of being hired. *British Journal of Management*, 27(4), 727-739.

- Kirkpatrick, L. A. (1992). An attachment-theory approach psychology of religion. *The International Journal for the Psychology of Religion*, 2(1), 3-28.
- Li, J. J., Ju, W., & Reeves, B. (2017). Touching a mechanical body: Tactile contact with body parts of a humanoid robot is physiologically arousing. *Journal of Human-Robot Interaction*, 6(3), 118–130.
- Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung [The innate forms of potential experience]. *Zeitschrift für Tierpsychologie*, 5, 233–519.
- Marchetti-Bowick, M. (2009). Is your roomba male or female? The role of gender stereotypes and cultural norms in robot design. *Intersect: The Stanford Journal of Science, Technology, and Society*, 2(1), 90-103.
- McGarty, C. (2018). Social categorization. In *Oxford Research Encyclopedia of Psychology*.
- Mieczkowski, H., Liu, S. X., Hancock, J., & Reeves, B. (2019, March). Helping not hurting: Applying the stereotype content model and bias map to social robotics. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 222-229). IEEE.
- Nass, C. (2004). Etiquette equality: Exhibitions and expectations of computer politeness. *Communications of the ACM*, 47(4), 35–37.
- Nass, C., & Lee, K. M. (2001). Does computer-synthesised speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, 7(3), 171–181.
- Parlangeli, O., Palmitesta, P., Bracci, M., Marchigiani, E., & Guidi, S. (2023). Gender role stereotypes at work in humanoid robots. *Behaviour & Information Technology*, 42(3), 316-327.
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., ... & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, 394(6696), 884-887.
- Perugia, G., Boor, L., van der Bij, L., Rikmenspoel, O., Foppen, R., & Guidi, S. (2023, March). Models of (Often) Ambivalent Robot Stereotypes: Content, Structure, and Predictors of Robots' Age and Gender Stereotypes. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 428-436).
- Perugia, G., Guidi, S., Bicchi, M., & Parlangeli, O. (2022, March). The shape of our bias: Perceived age and gender in the humanoid robots of the abot database. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 110-119). IEEE.

- Perugia, G., & Lisy, D. (2022). Robot's Gendering Trouble: A Scoping Review of Gendering Humanoid Robots and its Effects on HRI. *arXiv preprint arXiv:2207.01130*.
- Phillips, E., Zhao, X., Ullman, D., & Malle, B. F. (2018, February). What is human-like? decomposing robots' human-like appearance using the anthropomorphic robot (abot) database. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction* (pp. 105-113).
- Powers, A., & Kiesler, S. (2006, March). The advisor robot: tracing people's mental model from a robot's physical attributes. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-Robot Interaction* (pp. 218-225).
- Rabb, N., Law, T., Chita-Tegmark, M., & Scheutz, M. (2022). An attachment framework for human-robot interaction. *International journal of social robotics*, 1-21.
- Reeves, B., Hancock, J., & Liu, X. (2020). Social robots are like real people: First impressions, attributes, and stereotyping of social robots. *Technology, Mind, and Behavior*, 1(1), 1-37.
- Rollero, C., Peter, G., & Tartaglia, S. (2014). Psychometric properties of short versions of the ambivalent sexism inventory and ambivalence toward men inventory. *TPM. Testing, Psychometrics, Methodology in Applied Psychology*, 21(2), 149-159.
- Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception*, 38(8), 1211-1219.
- Savela, N., Turja, T., & Oksanen, A. (2018). Social Acceptance of Robots in Different Occupational Fields: A Systematic Literature Review. *International Journal of Social Robotics*, 10(4), 493–502.
- Sparrow, R. (2021). What robots represent... and why it matters. *Culturally sustainable social robotics: Proceedings of Robophilosophy*, 335, 5.
- Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3), 219-232.
- You, Z. J., Shen, C. Y., Chang, C. W., Liu, B. J., & Chen, G. D. (2006). A robot as a teaching assistant in an english class. *Proceedings - Sixth International Conference on Advanced Learning Technologies, ICAALT 2006*, 2006, 87–91.
- Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media and Society*, 8(3), 401–419.

Appendix

Appendix A - Participants nationality

Table 30

Nationality of participants in the study.

Nationality	Number of participants
South Africa	151
Portugal	115
Poland	78
United Kingdom	70
Italy	48
Mexico	35
United States	32
Greece	26
Hungary	14
Spain	13
Zimbabwe	10
Canada	10
Chile	8
Germany	8
Nigeria	7
Ireland	6
Netherlands	6
Slovenia	5
Latvia	5

Brazil	4
Turkey	3
Russian Federation	3
France	3
Czech Republic	3
Switzerland	2
Sweden	2
Finland	2
China	2
Belgium	2
Austria	2
Guatemala	2
Angola	1
Argentina	1
Australia	1
Bangladesh	1
Cameroon	1
Egypt	1
El Salvador	1
Estonia	1
Georgia	1
India	1
Israel	1
Kyrgyzstan	1
New Zealand	1

Palestinian Territory	1
Philippines	1
Syrian Arab Republic	1
Uganda	1
Bolivarian Republic of Venezuela	1
Vietnam	1

Note: Total number of participants included here is 696, as some participants' nationality was not recorded.

Appendix B - Survey

Age

Please, rate the following question by moving the slider [slider from 1 – 100]

1. In which moment of life would you place the robot in the image considering that 1 is like a newborn child and 100 like a very old person?

Gender

Please, rate your level of agreement with the following question [7-point Likert scale 1= completely disagree to 7 = completely agree].

1. How would you describe the robot in the image?
 - a. Feminine
 - b. Masculine
 - c. Gender neutral

Communion

Please, rate your level of agreement with the following statements when looking at the robot in the image [7-point Likert scale 1= strongly disagree to 7 = strongly agree].

1. This robot is affectionate.
2. This robot is compassionate.
3. This robot is tender.
4. This robot is gentle.
5. This robot is sympathetic.

Agency

Please, rate your level of agreement with the following statements when looking at the robot in the image [7-point Likert scale 1= strongly disagree to 7 = strongly agree].

1. This robot is able to defend its own beliefs.
2. This robot is willing to take a stand.
3. This robot has leadership abilities.
4. This robot is able to make decisions easily.
5. This robot has a strong personality.

Emotional reactions

Please, rate your level of agreement with the following questions when looking at the robot in the image [7-point Likert scale 1= not at all to 7 = extremely].

1. To what extent do you feel envy towards this robot?
2. To what extent do you feel jealousy towards this robot?
3. To what extent do you feel admiration towards this robot?
4. To what extent do you feel pride towards this robot?
5. To what extent do you feel pity towards this robot?
6. To what extent do you feel sympathy towards this robot?
7. To what extent do you feel contempt towards this robot?
8. To what extent do you feel disgust towards this robot?

Behavioural tendencies

Please, rate your level of agreement with the following questions when looking at the robot in the image [7-point Likert scale 1= not at all to 7 = extremely].

1. How likely would you be to help this robot?
2. How likely would you be to protect this robot?
3. How likely would you be to fight this robot?
4. How likely would you be to attack this robot?

5. How likely would you be to cooperate with this robot?
6. How likely would you be to associate with this robot?
7. How likely would you be to exclude this robot?
8. How likely would you be to demean this robot?

Quality of appearance cues

1. Does the robot have visible head hair?
 - a. *Yes*
 - i. Is the hair/ hair module:
 1. *Long*
 2. *Short*
 3. *Metallic*
 4. *Realistic (life-like, human-like)*
 5. *Blond/ Yellow*
 6. *Dark*
 7. *Grey White*
 8. *Other colour [blank field]*
 - b. *No*
2. Does the robot have a visible head?
 - a. *Yes*
 - i. Is the head:
 1. *Big (compared to shoulder's width and body size)*
 2. *Small (compared to shoulder's width and body size)*
 3. *Elongated*
 4. *Short*
 5. *Rounded*
 6. *Squared*
 7. *Spherical*
 8. *Oval*
 9. *Realistic (life-like, human-like)*

10. Mechanical

11. A screen or a tablet

12. A camera

b. No

3. Does the robot have a visible face?

a. Yes

i. Is the face:

1. Realistic (life-like, human-like)

2. Metallic

3. Mechanical

4. Chubby

5. Sunken

6. Expressive

7. Wrinkled

8. Bearded/ with moustache

9. With make-up

b. No

4. Does the robot have visible eyebrows?

a. Yes

i. Are the eyebrows:

1. Thick

2. Thin

3. Metallic

4. Realistic (life-like, human-like)

5. Straight

6. Curved

7. Raised

8. Frowning

b. No

5. Does the robot have visible eyelashes?

a. Yes

i. Are the eyelashes:

1. *Long*

2. *Short*

b. *No*

6. Does the robot have visible eyes?

a. *Yes*

i. Are the eyes:

1. *Big (compared to the head size)*

2. *Small (compared to the head size)*

3. *Round*

4. *Elongated*

5. *Expressive*

6. *Movable*

7. *Metallic*

8. *Realistic (life-like, human-like)*

9. *Mechanical*

10. *With big pupils*

b. *No*

7. Does the robot have a visible nose?

a. *Yes*

i. Is the nose:

1. *Long*

2. *Short*

3. *Rounded*

4. *Pointed*

5. *Narrow*

6. *Wide*

b. *No*

8. Does the robot have a visible mouth?

a. *Yes*

i. Is the mouth:

1. *Big*

2. *Small*

3. *Squared*
4. *Smiling*
5. *Minimalistic*
6. *Moveable*
7. *Red*
8. *Featuring full lips*
9. *Featuring thin lips*

b. *No*

9. Does the robot have a visible torso?

a. *Yes*

i. Is the torso:

1. *Long*
2. *Short*
3. *Metallic*
4. *Realistic (life-like, human-like)*
5. *Muscular*
6. *Slender*
7. *Chubby*
8. *Curvy*
9. *With mechanics visible*

ii. Does the torso have:

1. *Hourglass figure (a small waist compared to hips and shoulders)*
2. *Protruding chest (breast-like shape)*
3. *Neither is applicable*

b. *No*

10. Does the robot have visible arms?

a. *Yes*

i. Are the arms:

1. *Long*
2. *Short*
3. *Metallic*
4. *Realistic (life-like, human-like)*

5. *Muscular*
6. *Slender*
7. *Chubby*
8. *With mechanics visible*

b. *No*

11. Does the robot have visible hands?

a. *Yes*

- i. Are the hands:
 1. *Movable*
 2. *Able to grasp objects*
- ii. Does the robot have visible fingers?
 1. *Yes*
 - a. Are the fingers:
 - i. *Long*
 - ii. *Short*
 - iii. *Chubby*
 2. *No*

b. *No*

12. Does the robot have visible legs?

a. *Yes*

- i. Are the legs:
 1. *Long*
 2. *Short*
 3. *Metallic*
 4. *Realistic (life-like, human-like)*
 5. *Muscular*
 6. *Slender*
 7. *Chubby*
 8. *With mechanics visible*

b. *No*

13. Does the robot have visible skin?

a. *Yes*

- i. Is the skin:
 1. *Metallic*
 2. *Realistic (life-like, human-like)*
 3. *Black*
 4. *Brown*
 5. *White*
 6. *Pink*
 7. *Yellow*
 8. *Other colour* [blank field]

b. *No*

14. Does the robot have visible apparel (clothes)/ an apparel-like shape?

a. *Yes*

- i. Is the apparel:
 1. *T-shirt*
 2. *Skirt*
 3. *Trousers*
 4. *Dress*
 5. *Jacket*
 6. *Suit*
 7. *Mask*
 8. *Helmet*
 9. *Armor*
 10. *Tight*
 11. *Loose*
 12. *Formal*
 13. *Informal*
 14. *Pink*
 15. *Blue*
 16. *Dark*
 17. *White*
 18. *Other colour* [blank field]

b. *No*

Demographics

Age

Please write your answer here: [in numbers]

Gender

1. Woman
2. Man
3. Non-binary
4. I'd rather not say
5. I'd rather specify [blank field]

Familiarity with Artificial Intelligence (AI), robots, and science fiction

Rate your degree of familiarity with: [1= not familiar at all, 7= very familiar]

1. AI
2. Robots
3. Science fiction

Anthropomorphism tendency

Please, rate the extent to which you agree with the following statements [11-point Likert scale 0= not at all vs. 10=very much].

1. To what extent does the average robot have consciousness?
2. To what extent does the average computer have a mind of its own?
3. To what extent does technology—devices and machines for manufacturing, entertainment, and productive processes (e.g., cars, computers, television sets)—have intentions?

Ambivalent Sexism Inventory

Please, rate the extent to which you agree with the following statements [7-point Likert scale 1= strongly disagree vs. 7=strongly agree; B= Benevolent Sexism; H=Hostile Sexism].

1. Many women have a quality of purity that few men possess (B).

2. Women should be cherished and protected by men (B).
3. Women seek to gain power by getting control over men (H).
4. Every man ought to have a woman whom he adores (B).
5. Men are incomplete without women (B).
6. Women exaggerate problems they have at work (H).
7. Once a woman gets a man to commit to her, she usually tries to put him on a tight leash (H).
8. When women lose to men in a fair competition, they typically complain about being discriminated against (H).
9. Many women get a kick out of teasing men by seeming sexually available and then refusing male advances (H).
10. Women, compared to men, tend to have a superior moral sensibility (B).
11. Men should be willing to sacrifice their own well being in order to provide financially for the women in their lives (B).
12. Feminists are making unreasonable demands of men (H).

Appendix C - Pre-analysis: Appearance qualities that do not have a significant effect

Table 31

Appearance cues that do not have a significant effect on masculinity.

Appearance cue	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Short hair	-0.307	0.189	-1.62	-0.058	0.105
The presence of a torso	-0.199	0.260	-0.77	-0.046	0.443
Long torso	-0.039	0.193	-0.20	-0.008	0.838
Short torso	0.074	0.199	0.37	0.014	0.710
Metallic torso	0.156	0.195	0.80	0.040	0.424
Human-like torso	0.365	0.255	1.43	0.058	0.154
Curvy torso	-0.467	0.253	-1.84	-0.060	0.066
Long legs	0.048	0.251	0.19	0.009	0.849
Short legs	-0.102	0.251	-0.41	-0.020	0.683
Metallic legs	-0.169	0.241	-0.70	-0.042	0.483
Human-like legs	-0.154	0.401	-0.38	-0.016	0.701
Muscular legs	0.262	0.305	0.86	0.032	0.391
Slender legs	0.116	0.292	0.40	0.016	0.691
Chubby legs	0.191	0.251	0.76	0.028	0.447
Legs with visible mechanics	-0.027	0.227	-0.12	-0.006	0.903
Torso with hourglass figure	0.111	0.166	0.67	0.024	0.504

Note: Regression model: $R^2 = 0.123$, $F(22, 691) = 5.9$, $p < 0.0001$.

Table 32*Appearance cues that do not have a significant effect on femininity.*

Appearance cue	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Long hair	0.332	0.328	1.01	0.042	0.313
Long eyelashes	0.034	0.509	0.07	0.003	0.946
Full lips	0.016	0.220	0.07	0.002	0.941
White skin	-0.060	0.161	-0.37	-0.012	0.709
The presence of apparel	0.419	0.242	1.73	0.102	0.084
T-shirt	0.287	0.227	1.26	0.043	0.207
Jacket	-0.228	0.379	-0.60	-0.020	0.548
Suit	-0.606	0.435	-1.39	-0.050	0.164
Mask	-0.359	0.658	-0.55	-0.014	0.585
Helmet	-0.417	0.272	-1.53	-0.051	0.127
Armor	0.222	0.280	0.79	0.026	0.428
Tight	0.147	0.261	0.57	0.017	0.572
Loose apparel	0.023	0.255	0.09	0.003	0.926
Formal apparel	0.255	0.268	0.95	0.034	0.341
Informal apparel	-0.216	0.242	-0.89	-0.034	0.373
White apparel	0.265	0.204	1.30	0.049	0.194
Torso with protruding chest	0.092	0.181	0.51	0.017	0.610

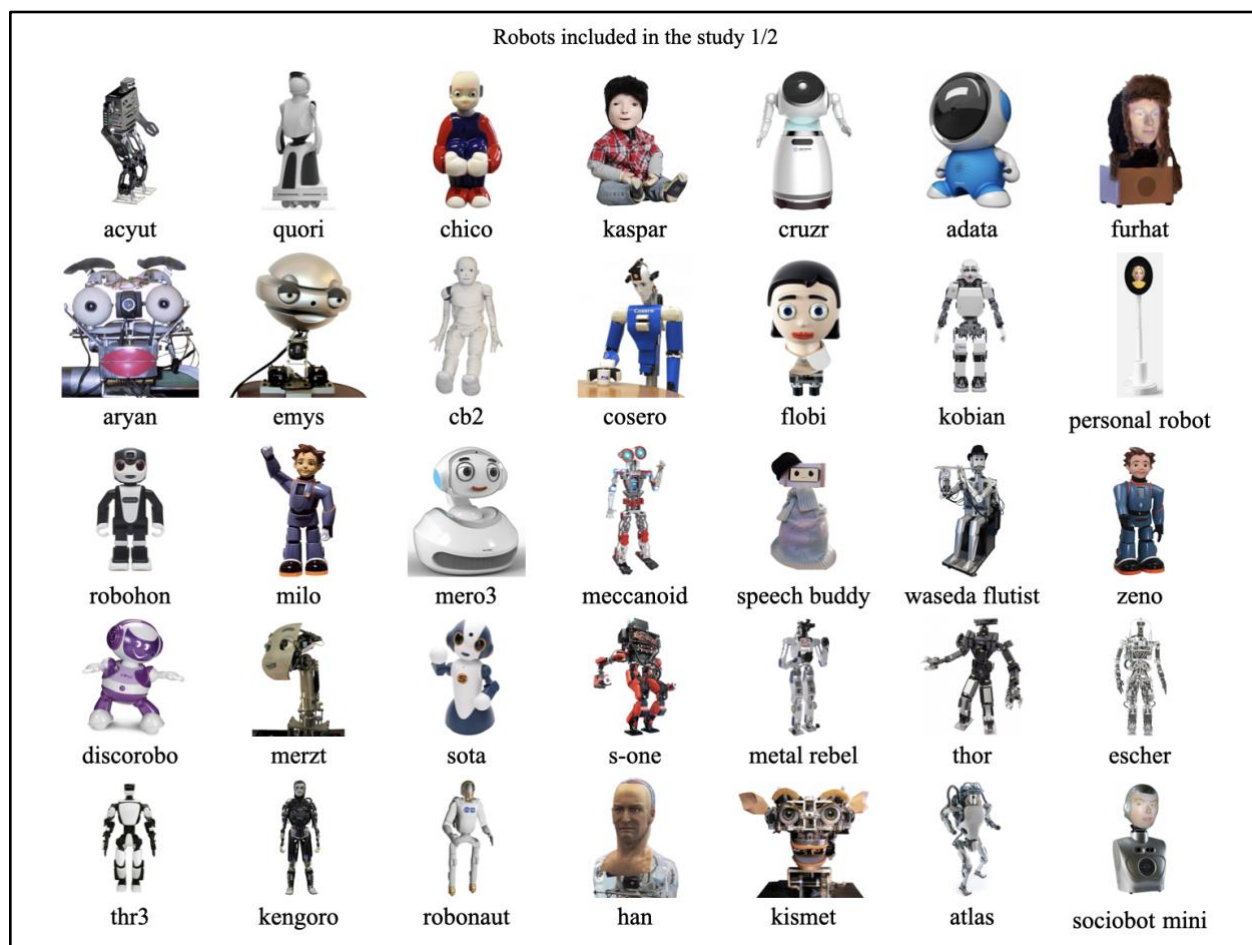
Note: Regression model: $R^2 = 0.283$, $F(24, 690) = 16$, $p < 0.0001$.

Table 32*Appearance cues that do not have a significant effect on age.*

Appearance cue	Unstandardized coefficients (B)	Robust standard error (SE)	t	Standardised coefficients (β)	Significance level (p)
Big head	-2.67	1.72	-1.55	-0.055	0.122
Big eyes	-0.289	2.62	-0.11	-0.004	0.912
Pointed nose	-4.526	2.39	-1.89	-0.079	0.059
Narrow nose	1.347	2.45	0.55	0.023	0.583
Wide nose	1.674	2.948	0.57	0.021	0.570

Note: Regression model: $R^2 = 0.115$, $F(10, 703) = 9.98$, $p < 0.0001$.

Appendix D - Robots included in this study



Robots included in the study 2/2



alter



felix



asimo



surena 3



nao



furo



ira



aimec 4



roboy



kibo



ibuki



albert einstein hubo



poppy



pino



robothespian



bina48



sophia



robina



aila



jia jia



kodomoroid



otonaroid



erica



nadine



gazeroid



gemini



posy



murata girl



sanbot



e-nuvo



mobiserv



talk torque 2



snackbot



ipal



padbot 3