Investigating the Emotional Impact of Social Robots

A Comparative Study on the Influence of Appearance and Application Area on Human Emotions

Tyra Wallén

Supervisor: Sam Thellman
Examiner: Tom Ziemke
Copyright

The publishers will keep this document online on the Internet – or its possible replacement – for a period of 25 years starting from the date of publication barring exceptional circumstances. The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/her own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security, and accessibility. According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement. For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its www home page: https://ep.liu.se/.

© 2023 Tyra Wallén
This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
Abstract

The rapid development of social robots, designed to interact with humans, has led to increased research on user acceptance and emotions in human-robot interaction. Social acceptance is an important area to investigate if the development of social robots is to be useful. Investigating how people feel about social robots is one tool to assess acceptance toward them, and research has shown that positive emotions could invoke higher acceptance. Possible factors that have been shown to affect peoples’ attitudes regarding social robots is (1) the human-likeness and appearance of the robot and (2) the application area of the robot. Therefore, this thesis research questions address the effect of human-likeness and application areas of social robots on people's emotions. The findings indicate that in the context of companionship, people have varying emotional responses based on the appearance of the social robot. Highly human-like robots evoke more positive emotions, while low human-likeness robots elicit more negative emotions. This suggests that individuals prefer human-like social robots in intimate interactions like companionship. The results also reveal an effect of application areas, where people respond more positively to highly human-like robots used for tasks like lecturing students or companionship for older adults. Regarding less human-like social robots, people tend to respond with greater positive emotions when used within commerce. This suggests that a simpler-looking robot with low human-likeness is more suitable for commercial applications. Negative emotions expressed in the healthcare condition may reflect mistrust in robots' abilities and the sensitivity of the healthcare area. Developers and designers should consider the emotional responses that might be evoked by the task or appearance of the social robot, to ensure successful integration into society.

*Keywords:* social robots, human-likeness, appearance, application area, HRI, emotions, acceptance
Acknowledgement

Biggest thanks to my supervisor Sam Thellman for helping me clarify my thoughts and providing me with valuable tips on how to proceed with my work. I also appreciate his patience in handling approximately 100 emails per day. Additionally, I would like to thank my seminar group for their input, as well as all my dear friends who may not have understood my work but still did their best to discuss ideas with me. I also want to express gratitude for the sixth floor of the studenthuset for providing an excellent study environment. Lastly, I thank myself.

Linköping, May 2023

Tyra Wallén
Table of Contents

1. Introduction
   1.1 Aim & Research Questions

2. Background
   2.1 The Social Robot
   2.2 Appearance of Social Robots
   2.3 Application Areas of Social Robots
   2.4 Emotional Responses Toward Social Robots
   2.5 The Affective Slider

3. Method
   3.1 Participants
   3.2 Stimuli
   3.3 Measures
   3.4 Procedure
   3.5 Pilot Study

4. Results
   4.1 What Degree of Human-Likeness Do the Different Social Robots Have?
   4.2 What Effect Does the Human-Likeness and Application Area of Social Robots Have on People’s Emotions (Pleasure)
      4.2.1 How Do People Feel About Social Robots With Different Degrees of Human-Likeness?
      4.2.2 How Do People Feel About Social Robots in Different Application Areas?
   4.3 What Effect Does the Human-Likeness and Application Areas of Social Robots Have on People’s Emotions (Arousal)
      4.3.1 How Do People Feel About Social Robots With Different Degrees of Human-Likeness?
      4.3.2 How Do People Feel About Social Robots in Different Application Areas?

5. Discussion
   5.1 What Degree Of Human-Likeness Do The Different Social Robots Have?
   5.2 What Effect Does the Human-Likeness and Application Areas of Social Robots Have on People’s Emotions
      5.2.1 How Do People Feel About Social Robots With Different Degrees of Human-Likeness?
      5.2.2 How Do People Feel About Social Robots in Different Application Areas?
   5.3 Emotions Towards the Social Robots and Their Application Area Based on Barrett & Russell’s Semantic Structure of Affect
      5.3.1 Companionship: Otonaroid vs. Padbot
      5.3.2 Otonaroid: Companionship vs. Healthcare
      5.3.3 Padbot: commerce vs. Companionship
      5.3.4 Nao and Pepper: Commerce vs. Healthcare
   5.4 Method
   5.5 Future Studies

6. Conclusions

References
1. Introduction

Over the past few years, the development and research of social robots, i.e., robots that are designed to interact with people (Fong et al., 2003), has rapidly increased in various fields, including healthcare, education, companionship (Naneva et al., 2020) and commerce (Niemelä et al., 2019). In line with this, research on user acceptance towards social robots has also increased in order to map out which potential factors in human-robot interaction (HRI) might be worth considering when developing social robots (Lehmann et al., 2020; Liu et al, 2022; Naneva et al., 2020; Tay et al., 2014). Acceptance can be defined as the intention to use or interact with a social robot (Lehmann et al., 2020), and is an important area to investigate if the development of social robots is to be useful. Social robots have been developed with differing human-likeness, and some research has been done on the relationship between the human-likeness of robots and people’s acceptance toward them. While studies have shown that robots with higher human-likeness elicit more positive responses from individuals, robots that are too human-like can be experienced as uncanny (Mori et al., 2012). The application area of the robot also plays a crucial role in determining people's attitudes towards social robots, which will be further explained in the background section of this thesis. Investigating how people feel about social robots is one tool to assess acceptance toward them (Nomura et al., 2008), and research has shown that positive emotions could invoke higher acceptance (Hwang et al., 2013).

1.2 Aim & Research Questions

The increased development of social robots in various fields has generated significant interest in their potential applications, and understanding how people perceive and respond to these robots is crucial for their successful integration into society. The current study therefore aims to provide a better understanding of how the human-likeness and application area of social robots contribute to people’s emotions (and thereby acceptance) toward social robots. To meet this aim, the study addresses the following research questions (RQs):

1. What effect does the human-likeness and application areas of social robots have on people’s emotions?
1a. How do people feel about social robots with different degrees of human-likeness?
1b. How do people feel about social robots in different application areas?
2. Background

The background of the current study will contain previous research about human-likeness of social robots, various application areas, emotions toward social robots and the use of the Affective Slider as a way of measuring emotions. The background will also clarify the relevance of examining peoples emotions towards social robots.

2.1 The Social Robot

Within the science of social robotics, there is no clear universal definition of a social robot (Henschel et al., 2021). How a social robot behaves and what actually makes it social, seems to vary from robot to robot. However, one common factor among social robots is the ability to interact more or less with human beings. The role of social robots as social partners has led to the increasing application and research of human-robots interaction (HRI) within various fields, such as in education, care, and companionship (Mejia & Kajikawa, 2017). Just as the diversity in capabilities among social robots, they also do not always have a homogenous appearance where robots could be illustrated as a pet or android (human-like) etc. (Lehmann et al., 2020). Due to the current state in the development of technology, social robots still have their limitations. Although research suggests they have growing potential in improving human life (Henschel et al., 2021), which makes it immensely crucial to examine the factors that make the human more or less comfortable interacting with social robots.

2.2 Appearance of Social Robots

The appearance of social robots can influence the level of acceptance among potential users (Lehmann et al., 2020), and is therefore an important factor to consider when developing HRI. If the user experiences positive emotions towards the appearance, it could invoke higher acceptance (Hwang et al., 2013). Different categories of appearances have been discussed (functional, cartoon-like, functional, anthropomorphic, etc.) where anthropomorphic robots in social HRI seem to be perceived with greater acceptance (Lehmann et al., 2020). Since user acceptance is a crucial element for successful and intuitive HRI development, there has been a growing trend to create social robots capable of mimicking both human behavior and appearance (Tay et al., 2014). One such example is the humanoid social robot Nadine, which is modeled after a human (MIRALab, n.d.).

However, there is a fine line between anthropomorphic robots and too much human-likeness, where robots that resemble a human too much (but not perfectly) could trigger more negative feelings and less acceptance. Several studies have investigated the effects of different degrees of human-likeness, referring to the hypothesis “Uncanny Valley” (Mara et al., 2022; Lehmann et al., 2020; Laakasuo et al., 2021). This hypothesis suggests that a feeling of uncanniness occurs when people perceive an artificial agent as too human-like (Mori et al., 2012). Mara et al. (2022) examined participants’ likeability of a robot and the perceived anthropomorphism. Robots with low to medium human-likeness characteristics correlated with higher likeness (more positive responses), in accordance with the hypothesis of “Uncanny Valley”. Although, this study did not draw any conclusions regarding responses towards higher levels of
anthropomorphism due to sparse utilization of these robots in previous studies. This study aims to use highly human-like social robots to fill the gap.

2.3 Application Areas of Social Robots

Apart from appearances, the intention to use social robots varies to some extent depending on the situation in which HRI occurs. A study showed that people more easily accepted social robots having an occupation that was in accordance with their stereotypes (Tay et al., 2014). In an additional study conducted on older adults’ attitudes and emotions towards robots, the participants expressed a more negative attitude regarding robots in care situations (Lehmann et al., 2020). The current study aims to examine the use of robots in four broad categories of application areas, partly inspired from Naneva et al. (2020): healthcare, education, companionship and commerce.

Research about the integration of social robots in healthcare has been an increasingly popular area of interest. A review and survey study exploring attitudes of healthcare professionals towards the use of social robots, found overall positive opinions regarding the matter (Vänni & Salin, 2019). This may be due to the many benefits of social robots in healthcare that the healthcare workers and educators expressed. Robots in the healthcare sector could save time and increase meaningfulness and productivity by assisting the workers. The areas in which the robot could assist the most were reported to be indirect nursing care, such as cleaning and managing food, followed by direct nursing care related to patient safety, such as monitoring, giving reminders, and raising alarms.

Education is another area identified by Naneva et al. (2020), where social robots could be used to assist with teaching and social interaction among students. A review study conducted by Belpaeme et al. (2018) found positive learning outcomes of the use of social robots in education, which were similar to human tutoring on restricted tasks. These included increasing cognitive outcomes, such as knowledge and comprehension in learners, and increasing affective outcomes, such as the learner being receptive, responsive or attentive. The studies reviewed used a wide range of social robots for education; however, Belpaeme et al. (2018) were unable to compare the outcomes of the different robots due to the varying methodologies employed. Nonetheless, it is noteworthy that Nao was reported to be the most popular robot used in education. Demographic variables have been shown to influence people's attitudes toward the use of social robots within education, with less favourable perceptions among older adults (aged >55) and those with little to no experience with robots (Smakman et al., 2021).

Companionship was the third application area inspired from Naneva et al. (2020). The article defined companionship robots as having the purpose of interacting socially with humans over an extended period of time and providing companionship. The present study focuses on the application of social robots in companionship, specifically in relation to older adults. A systematic review study conducted on social robots for older adults revealed that social robots may have a favorable impact on a variety of factors, for example stress and loneliness (Pu et
al., 2019). Older adults' emotions and attitudes towards social robots have been shown to depend on the appearance of the robot (Lehmann et al., 2020).

Social robots have also been studied in the realm of commerce, such as in the context of shopping malls, where the robot had tasks such as providing directions, greeting and welcoming, carrying bags, and engaging in pleasant conversations with customers (Niemelä et al., 2019). In the study conducted by Niemelä et al. (2019), the social robot Pepper was deployed, and the findings indicate a generally positive attitude, with high levels of acceptance towards its usage.

Based on this brief review of previous research stated above, a generally positive attitude towards robots in these specific domains can be found with some differences depending on appearance and demographics. However, the literature provides limited insights about the influence of robot appearance on individuals' experiences of the use of social robots in the different situations. Therefore, the current study will apply a combination of the robot’s appearance and application area to investigate people’s emotions.

2.4 Emotional Responses Toward Social Robots

External stimuli from the perceived environment can influence human emotional responses in the form of pleasure and arousal (Mehrabian & Russell, 1974). If social robots are to enter our everyday lives, it is crucial to understand their potential impact on human emotions, as it could affect human behavior, for example the willingness to interact with a robot (Nomura et al., 2008). Nomura et al. (2008) examined people’s emotions to predict human behavior in human-robot interaction. The researchers found a relationship between negative emotions and the tendency to avoid communicating with a robot. If the development of social robots is to progress in a positive direction, the users have to be onboard. Therefore, it is crucial to further investigate the specific elements of robots that elicit certain emotions.

Jung et al. (2021) conducted a study exploring affective (favorability) and cognitive (trust) responses to robots with differing human-likeness in two service contexts (hotel reception and tutoring). To do this, the researchers measured the participants favorability for each robot and the participants trust toward the robot being used as a hotel reception staff or tutor. The study found a significant effect of affective response on trust, implying that higher favorability could elicit higher trust. The analysis also revealed that the high-expertise humanoid (used for tutoring) generated more positive affective and cognitive responses compared to the low-expertise humanoid (used for hotel reception), for all robots. The most human-like robot scored highest on favorability and trust for both hotel and tutoring, and the least human-like robot (Pepper) scored second highest on favorability for both hotel and tutoring. The results from the researchers study suggests that people's evaluations of robots performing specific tasks, depends on the nature of the task as well as the appearance of the robot.
2.5 The Affective Slider

The current study aims to measure the participants' emotions towards the use of different social robots in various application areas. One approach to do this is by using the affective slider (AS), a digital self-reporting tool designed to measure human emotions (Betella & Verschure, 2016). The AS tool consists of two sliders, one measuring pleasure and one measuring arousal. The method is a simplified version of the Self-Assessment Manikin (SAM), which measures pleasure, arousal and dominance. Betella and Verschure (2016) demonstrated outcomes similar to those obtained from the SAM approach, with the added benefit of not requiring written instructions and being easily accessible via modern devices, such as tablets and smartphones.

The two continuous scales in the AS are located on top of each other in a randomized order over the trials. Two emotional faces are illustrated at each side of the slider, with a monochromatic neutral color scheme: unhappy and happy faces for pleasure, sleepy and wide-awake for arousal (figure 1). When using the AS in an experiment, a stimulus is presented and the participants are asked to rate how they feel about the stimuli using both sliders.

![Figure 1. The “Affective Slider” (AS), arousal (top) and pleasure (bottom) (Betella & Verschure, 2016).](image)

A way of analyzing results from the AS is by using Barrett and Russell’s (1998) model of emotion. The authors refer to pleasure as valence, and arousal as activation. They argue that these two dimensions are largely independent of each other and can be illustrated on a two-dimensional affective model (figure 2).
Figure 2. Barrett and Russell's (1998) semantic structure of affect. The letters x and y represent semantic components: x = pleasantness; y = activation.
3. Method

To answer the research questions, a quantitative research method was employed. A survey was designed with four different social robots and application areas. Thirty-six participants were recruited through Prolific.

3.1 Participants

After conducting a G*Power analysis of the current study’s sample size, it was determined that thirty-six participants were needed. The participants were recruited using Prolific (n.d.) . Prolific is a crowdsourcing marketplace that can be used for collecting high quality data for surveys. Given the study’s desire to obtain results from a broad population, Prolific was a suitable tool for a representative sample size. Requirements were established for survey participants, including a minimum approval rate of 95%, a minimum of 100 previous submissions, and location restricted to the United States. These requirements have been employed in previous studies on prolific and provided high quality data (Peer et al., 2022). A total of 37 participants took part in the study (15 men, 21 women and 1 non-binary) and the ages varied between 22 and 72 years old ($M = 41.86, SD = 14.57$). 44.6% had a high school degree, 36.8% had a bachelor’s degree, 7.9% had a master’s degree, 2.6% had a doctorate degree, and 5.3% had none of these. 39% had not seen a social robot before, 23.7 % had seen a social robot before, and 34.2% were unsure. 47.4% had not interacted with a social robot before, 21.1% had interacted with a social robot before, and 28.9% were unsure.

3.2 Stimuli

Each pictured stimulus was carefully picked from The ABOT Database (n.d.) to ensure that the robots had somewhat different characteristics and human-likeness score.

![Figure 3. Padbot (n.d.).](image1)

![Figure 4. Nao (n.d.).](image2)
The rationale behind selecting these stimuli was to enable the utilization of robots exhibiting varying degrees of human-likeness, which includes low, medium, and high levels. The ABOT Database (n.d.) has robots with human-likeness scores between 1 - 100. The scores are based on a systematic analysis of their collection comprising 251 robots, where their research identified three prominent categories of human-likeness, namely surface, face, and body manipulators. According to the database, Padbot’s score is 4.13 (figure 3), Nao’s score is 45.92 (figure 4), Otonaroid’s score is 88.74 (figure 5) and Pepper’s score is 42.17 (figure 6). Nao and Pepper are two prominent social robots within the research field (e.g. Thunberg et al., 2017; Manzi et al., 2021; Mubin et al., 2018). Therefore, despite their comparable scores in human-likeness, it is pertinent to investigate potential variances in public perception of these robots across different areas of application.

The chosen application areas were inspired from Naneva et al. (2020) who made a systematic review of trust, acceptance, attitudes and anxiety towards social robots. After analyzing ninety-seven studies with over 13,000 participants, they found several potential factors influencing people’s attitudes, including the robots’ domain of application. Six broad domains of application were identified in the study: companion robotics and domestic assistance, healthcare, education, HRI, general application and pediatric care. The current study adopted three of these: healthcare, education and companion robotics. In addition to the previously mentioned areas, the study encompasses the application area of commerce (see background). The application areas were formulated as follows (varying condition marked in bold):

**Healthcare**
Please rate how you feel about the pictured robot being used to **assist a doctor before, during and after surgery** using BOTH sliders below. (Don’t think too much about it, just rate how you feel).
Commerce
Please rate how you feel about the pictured robot being used to assist with navigation and payment when shopping using BOTH sliders below. (Don’t think too much about it, just rate how you feel).

Education
Please rate how you feel about the pictured robot being used to lecture students at school using BOTH sliders below. (Don’t think too much about it, just rate how you feel).

Companion robotics (older adults)
Please rate how you feel about the pictured robot being used to take care of and accompany older adults using BOTH sliders below. (Don’t think too much about it, just rate how you feel).

3.3 Measures
To assess the participants’ emotional responses towards social robots with varying degrees of human-likeness in different application areas, two measures were used in the survey. Firstly, to measure the level of human-likeness, the participants were asked to estimate the human-likeness for each robot (the questions were randomized):

Please rate how much this robot looks like a human using the slider below. (Figure 7).

![Figure 7. Scale for human-likeness.](image)

Thereafter, the study used sliders based on “the affective slider” to measure the participants emotional responses (pleasure and arousal) to each social robot and application area. Pleasure and arousal were measured separately on two different scales (figure 8, 9). The scales were continuous from 0 to 100, where 0 represented unpleasant/non-aroused and 100 represented pleasant/aroused. However, the numbers were not displayed in order to make them intuitive and reliable. In the survey, the participants were requested to evaluate how they feel about the use of the social robot in a particular area, based on the pleasure and arousal dimensions. The order in which pleasure and arousal were presented was randomized. The varying condition was marked in bold:

Please rate how you feel about the pictured robot being used to assist a doctor before, during and after surgery using BOTH sliders below. (Don’t think too much about it, just rate how you feel). (This is one of the four application area descriptions, see section 3.3).
3.4 Procedure

A quantitative online survey was constructed to reach a broad group of respondents. The surveys were created in the digital questionnaire tool SurveyMonkey (https://www.surveymonkey.com/), published on Prolific and consisted of 19 parts. Initially, the participants were presented with the purpose and structure of the study. To ensure the quality of the study, a definition of social robots was given, inspired by de Graaf et al. (2017):

*Social robots are autonomous robots that can operate independently in our everyday environments, interact more or less with humans and assist us in our everyday life. Social robots can interact with humans by speech, gestures and expressions in a humanlike way by following social behaviors and rules.*

After agreeing to participate in the study, the first section consisted of demographic data (age, gender, education) together with questions about previous experiences with social robots. The subsequent section aimed to investigate the degree to which the social robots (figure 3, 4, 5, 6) resemble humans. To examine the impact of the application area and appearance of each robot, the following sections consisted of a total of 16 randomized questions, in which each social robot was presented with each application area (4x4). On the top of the page was an image of a social robot, underneath it was the question about how the participants feel about

*11. Move the slider to rate your level of **Pleasure**

![Pleasure Scale](image)

*Figure 8. Scale for pleasure.*

*12. Move the slider to rate your level of **Arousal**

![Arousal Scale](image)

*Figure 9. Scale for arousal.*
the robot being used in a certain area, followed by the pleasure and arousal scales. The participants had to move the scales in order to move on to the next question.

3.5 Pilot Study

Prior to publishing the survey on Prolific, a pilot study was conducted to evaluate the design of the survey (i.e. validate the displayed questions and address any potential problems). Ten test participants completed the survey. All ten test subjects were students at Linköpings University, with six identifying as women and four identifying as men. On average, the survey was completed in 6.75 minutes. Participants were seated next to the researcher during the test and were encouraged to express their thoughts and share any uncertainties they had regarding the test. Another pilot study was conducted on five additional participants from Prolific, three men and two women, to validate the absence of any technical issues before publishing the final survey. The results from the pilot studies were not included in the analysis. However, a few improvements were made based on the feedback provided by the participants.
4. Results

To answer the research questions, two-way repeated measures ANOVA and simple main effects were run in SPSS to determine the effect of human-likeness and application areas of social robots on the participants’ emotions. The results for pleasure and arousal were analysed separately. For pictures of the robots and descriptions of the application areas, see section 3.2.

4.1 What Degree of Human-Likeness Do the Different Social Robots Have?

The purpose of analyzing the human-likeness of the different robots was to understand how it affects the participants emotions in various application areas. Otonaroid scored highest on human-likeness \((M = 82.59, SD = 16.19)\), followed by Nao \((M = 26.81, SD = 23.88)\) and Pepper \((M = 25.95, SD = 22.21)\), and Padbot scored lowest \((M = 3.11, SD = 5.84)\).

4.2 What Effect Does the Human-Likeness and Application Area of Social Robots Have on People’s Emotions (Pleasure)?

A two-way repeated measures ANOVA was run to determine the effect of human-likeness and application areas of social robots on the participants’ appreciated level of pleasure. Analysis of the studentized residuals showed that there were no outliers, as assessed by no studentized residuals greater than ±3 standard deviations. The answers on pleasure were normally distributed \((p > .05)\) except for eight variables \((p = .008 - .024)\), as assessed by the Shapiro-Wilk test for normality. However, skewness and kurtosis are not greater than ±2, indicating a somewhat normal distribution. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for the two-way interaction, \(\chi^2(2) = 70.5, p = .007\), therefore, a Greenhouse-Geisser correction was made. There was a statistically significant interaction between robot and application area on the pleasure dimension, \(F(6.201, 223.244) = 4.413, p < .001, \epsilon = .689\) (figure 10). Therefore, simple main effects were run.
4.2.1 How Do People Feel About Social Robots With Different Degrees of Human-Likeness?

Appreciated pleasure was not statistically significantly different between the robots in the healthcare area, $F(2.195, 79.015) = .405, p = .687$, in the education area, $F(2.528, 91) = 2.217, p = .112$, and in the commerce area, $F(2.143, 77.136) = 2.736, p = .067$. Appreciated pleasure was statistically significantly different between the robots in the companionship area, $F(2.187, 78.704) = 5.204, p = .006$.

Appreciated pleasure was higher for Otonaroid ($M = 52.62, SD = 31.605$) compared to Padbot ($M=35.32, SD = 29.319$) in the companionship area, a statistically significantly mean increase of $17.297, SE = 5.743, p = .028$. Appreciated pleasure was also higher for Pepper ($M = 49.35, SD = 28.753$) compared to Padbot ($M=35.32, SD = 29.319$) in the companionship area, a statistically significantly mean increase of $14.027, SE = 3.864, p = .005$.

The participants expressed significantly greater pleasure for the most human-like robot, Otonaroid, and the robot Pepper, compared to the least human-like robot, Padbot, when the robot was used to take care of and accompany older adults.

4.2.2 How Do People Feel About Social Robots in Different Application Areas?


---

**Figure 10.** Plot for the interaction between robot and application area on the pleasure dimension.
Otonaroid
Appreciated pleasure was higher in the education area \((M = 51.81, SD = 30.0)\) compared to the healthcare area \((M = 39.35, SD = 30.288)\), a statistically significantly mean increase of 12.459, \(SE = 3.786, p = .013\). Appreciated pleasure was also higher in the companionship area \((M = 52.62, SD = 31.605)\) compared to the healthcare area \((M = 39.35, SD = 30.288)\), a statistically significantly mean increase of 13.270, \(SE = 3.867, p = .009\).

For the most human-like robot, Otonaroid, the participants expressed significantly greater pleasure when the robot was used to lecture students at school and take care of and accompany older adults compared to when the robot was used to when the robot was used to assist a doctor before, during and after surgery.

Padbot
Appreciated pleasure was higher in the commerce area \((M = 53.49, SD = 26.130)\) compared to the healthcare area \((M = 39.35, SD = 28.695)\), a statistically significantly mean increase of 14.135, \(SE = 3.851, p = .005\). And compared to the education area \((M = 41.00, SD = 32.225)\), a statistically significantly mean increase of 12.486, \(SE = 3.849, p = .015\). And compared to the companionship area \((M = 35.32, SD = 29.319)\), a statistically significantly mean increase of 18.162, \(SE = 3.99, p < .001\).

For the least human-like robot, Padbot, the participants expressed significantly higher pleasure when the robot was used to assist with navigation and payment when shopping compared to all the other application areas.

Nao
Appreciated pleasure was higher in the commerce area \((M = 56.24, SD = 29.934)\) compared to the healthcare area \((M = 39.70, SD = 29.561)\), a statistically significantly mean increase of 16.541, \(SE = 4.708, p = .007\). And compared to the companionship area \((M = 44.05, SD = 28.747)\), a statistically significantly mean increase of 12.189, \(SE = 3.858, p = .019\).

For the robot Nao, the participants expressed significantly higher pleasure when the robot was used to assist with navigation and payment when shopping compared to when the robot was used to assist a doctor before, during and after surgery, and when the robot was used to take care of and accompany older adults.

Pepper
Appreciated pleasure was higher in the commerce area \((M = 58.30, SD = 26.372)\) compared to the healthcare area \((M = 42.97, SD = 30.124)\), a statistically significantly mean increase of 15.324, \(SE = 3.963, p = .003\).

For the robot Pepper, the participants expressed significantly higher pleasure when the robot was used to assist with navigation and payment when shopping compared to when the robot was used to assist a doctor before, during and after surgery.
4.3 What Effect Does the Human-Likeness and Application Areas of Social Robots Have on People’s Emotions (Arousal)?

A two-way repeated measures ANOVA was run to determine the effect of human-likeness and application areas of social robots on the participants’ appreciated level of arousal. Analysis of the studentized residuals showed that there were no outliers, as assessed by no studentized residuals greater than ± 3 standard deviations. The answers on pleasure were normally distributed ($p > .05$) except for five variables ($p = .003 - .035$), as assessed by the Kolmogorov-Smirnov test of normality on the studentized residuals. However, skewness and kurtosis are not greater than ± 2, indicating a somewhat normal distribution. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for the two-way interaction, $\chi^2(2) = 99.6, p < .001$, therefore, a correction was made using the Greenhouse-Geisser. There was a statistically significant interaction between robot and application area on the arousal dimension, $F(6.028, 217.009) = 2.176, p = .046, \varepsilon = .67$ (figure 11). Therefore, simple main effects were run.

![Estimated Marginal Means of MEASURE_1](image)

**Figure 11.** Plot for the interaction between robot and application area on the arousal dimension.

4.3.1 How Do People Feel About Social Robots With Different Degrees of Human-Likeness?

Appreciated arousal was not statistically significantly different between the robots in the healthcare area, $F(2.458, 88.5) = .302, p = .784$, or in the commerce area, $F(2.4, 86.478) = 2.538, p = .075$. However, appreciated arousal was statistically significantly different between the robots in the education area, $F(2.537, 91.32) = 4.606, p = .007$, and in the companionship area, $F(2.049, 73.775) = 5.562, p = .005$. 

17
Education
Appreciated arousal was lower for Padbot ($M = 33.49$, $SD = 31.549$) compared to Otonaroid ($M = 48.59$, $SD = 31.815$), a statistically significantly mean decrease of -15.108, $SE = 4.345$, $p = .008$.

The participants expressed significantly less arousal for the least human-like robot, Padbot, compared to the most human-like robot, Otonaroid, when the robot was used to lecture students at school.

Companionship
Appreciated arousal was lower for Padbot ($M = 30.89$, $SD = 29.799$) compared to Otonaroid ($M = 49.22$, $SD = 32.448$), a statistically significantly mean decrease of -18.324, $SE = 5.765$, $p = .018$. And compared to Pepper ($M = 41.03$, $SD = 30.550$), a statistically significantly mean decrease of -10.135, $SE = 3.597$, $p = .047$.

The participants expressed significantly lower arousal for the least human-like robot, Padbot, compared to the most human-like robot, Otonaroid, and the robot Pepper, when the robot was used to take care of and accompany older adults.

4.3.2 How Do People Feel About Social Robots in Different Application Areas?
Appreciated arousal was not statistically significantly different between the application areas for Padbot, $F(3, 108) = 2.025$, $p = .115$, or for Pepper, $F(2.52, 90.8) = 1.3$, $p = .283$. However, appreciated arousal was statistically significantly different between the application areas for Otonaroid, $F(2.5, 90.39) = 3.9$, $p = .016$, and for Nao, $F(3, 108) = 2.758$, $p = .046$.

Otonaroid
Appreciated arousal was lower within healthcare ($M = 37.92$, $SD = 32.056$) compared to Education ($M = 48.59$, $SD = 31.815$), a statistically significantly mean decrease of -10.676, $SE = 3.178$, $p = .011$. And compared to companionship ($M = 49.22$, $SD = 32.448$), a statistically significantly mean decrease of -11.297, $SE = 3.793$, $p = .031$.

The participants expressed significantly lower arousal when the most human-like robot, Otonaroid, was used to assist a doctor before, during and after surgery, compared to when it was used to lecture students at school and take care of and accompany older adults.

Nao
Appreciated arousal was lower within companionship ($M = 40.16$, $SD = 31.984$) compared to commerce ($M = 49.59$, $SD = 32.935$), a statistically significantly mean decrease of -9.432, $SE = 3.12$, $p = .027$.

The participants expressed significantly lower arousal when the robot Nao was used to take care of and accompany older adults, compared to when it was used to assist with navigation and payment when shopping.
5. Discussion

The aim of this thesis was to better understand the impact of social robots’ human-likeness and application area on people’s emotions (pleasure and arousal). If social robots are to successfully integrate into society, it is crucial to explore the factors influencing people’s acceptance. In this section, the results will be discussed based on the study's research questions, and some significant results will be discussed based on Barrett and Russell's (1998) semantic structure of affect.

5.1 What Degree Of Human-Likeness Do The Different Social Robots Have?

The robots in the study had different levels of human-likeness, the means of which were as follows: Otonaroid (\(M = 82.59, SD = 16.19\)), Padbot (\(M = 3.11, SD = 5.84\)), Nao (\(M = 26.81, SD = 23.88\)) and Pepper (\(M = 25.95, SD = 22.21\)). An objective was to find robots with low, medium and high human-likeness from the ABOT Database (n.d.), and the study successfully obtained robots with high and low levels of human-likeness. According to the ABOT Database, Nao’s human-likeness score is 45.92 and Pepper’s score is 42.17. However, the mean human-likeness score in the current study for Nao and Pepper was relatively low compared to those in the ABOT Database. Therefore, conclusions regarding social robots with a medium human-likeness score can not be made.

5.2 What Effect Does the Human-Likeness and Application Areas of Social Robots Have on People’s Emotions?

The results from the two-way repeated measures ANOVA showed a statistically significant interaction between robot and application area on both the arousal and pleasure dimension. This indicates that the social robot’s appearance and application area have an effect on the participant’s perceived pleasure and arousal. The finding is in accordance Mehrabian and Russell (1974) who argued that external stimuli can influence human emotional responses, and studies claiming there is a variation on people's responses towards social robots depending on the robot’s characteristics (Tay et al., 2014; Lehmann et al., 2020; Jung et al., 2021). The robot’s appearances had a varying degree of human-likeness, in which might be influencing the participants emotional responses (Mara et al., 2022; Lehmann et al., 2020; Laakasuo et al., 2021). To further investigate what effect the appearances (human-likeness) and application areas of social robots have on the participant’s emotions, the results from the simple main effects will be discussed.

5.2.1 How Do People Feel About Social Robots With Different Degrees of Human-Likeness?

The analysis revealed varying emotional responses within the context of companionship, indicating that people feel different about using a social robot as a companion among older adults depending on the appearance of the robot. However, within education, healthcare and commerce, the human-likeness or appearance of the social robot does not have an effect on people's emotions in the current study. The results suggest that developers and designers of
social robots as companions should consider how human-like features of the robot may evoke emotional responses, and how this may affect people's acceptance. Companionship robotics can be viewed as a quite intimate application area and raises some ethical concerns, such as older adults' privacy and attachment to robots (Ostrowski et al., 2022). Therefore, a social robot used to “take care of and accompany older adults” might be experienced as a sensitive application area for people, and hence the varying results depending on the appearances of the social robot. In the context of companionship, people seem to express greater positive emotions for highly human-like social robots (such as Otonaroid) or the robot Pepper, and more negative emotions for social robots with low human-likeness (with a design like Padbot). This may indicate that individuals would rather see a human-like robot keeping older adults company, than a mechanical robot that does not resemble a human. One could hypothesize that individuals are more likely to trust a robot with human-like features, especially in areas involving intimate interactions like companionship. This tendency might stem from beliefs that highly human-like robots are more competent in human-related tasks, as in accordance with the results from Jung et al. (2021).

It seems that individuals experience lower activation or stimulation when the least human-like robot was used to “lecture students at school”, and higher activation or stimulation for the most human-like social robot. One possibility is that participants may have found the social robot with low human-likeness less engaging or stimulating, resulting in lower arousal. On the other hand, the robot with high human-likeness might have elicited a stronger response or captured more attention, leading to higher arousal levels. However, since there was no significance on the pleasure dimension within the education condition, it is difficult to discuss how positive or negative the results are.

5.2.2 How Do People Feel About Social Robots in Different Application Areas?
Overall, the results indicate varying emotional responses depending on which area the social robot is used within. Based on this, it could be interpreted that people’s emotional responses are more sensitive towards the application area rather than the appearance or human-likeness of the robot. Firstly, people tend to respond with greater positive emotions regarding highly human-like social robots (such as Otonaroid), when it is used for tasks such as “lecture students at school” and “take care of and accompany older adults” compared to the task “assist a doctor before, during and after surgery”. A possible reason for this could be that the participants interpret highly human-like robots as more fitted for direct human contact that includes socializing, compared to more technical or clinical contexts. The negative expressed emotions towards using the most human-like robot within healthcare could possibly be due to the formulation of the healthcare condition, which stated that the robot would assist during surgery. Surgery might be a task that is taken very seriously, meanwhile if the robot had indirect nursing tasks within the healthcare such as cleaning or managing food, the attitudes could have been more positive. According to the review study by Vänni and Salin (2019), people have overall positive opinions regarding social robots helping out with indirect nursing tasks. Regarding the educational context, the positive emotional responses are in accordance with Jung et al. (2021), where the participants expressed high favorability and trust for the human-like robot used for tutoring. If developing and designing a highly human-like robot, as
Otonaroid, it would be less fitted for assisting during surgery and more suitable for companionship and lecture students at school, based on this study’s results.

A common theme for the other social robots with low human-likeness (Nao, Pepper and Padbot), are that people appear to experience them with greater positive emotions when used within commerce (“assist with navigation and payment when shopping”) compared to any other application area. These results are fairly expected, partly because of Pepper’s and Nao’s widespread usage within the research field (e.g. Thunberg et al., 2017; Manzi et al., 2021; Mubin et al., 2018). Perhaps people have seen social robots with similar appearances as Nao, Pepper and Padbot in a shopping mall context before (e.g. Niemelä et al., 2019) making it easier to accept and thereby having higher positive emotions regarding the condition. Since this trend in emotional responses only applies for the less human-like social robots (and not Otonaroid, who actually dropped in the commerce area), it suggests that a more simple looking social robot with somewhat low human-likeness would be more suited and positively experienced within commerce. The positive emotions expressed towards these robots in the commerce condition compared to the other areas, could indicate that people doubt the robots abilities to perform the other, more “advanced” tasks (such as lecture students at school), due to their mechanical looking features. Meanwhile, people’s tendency to express positive emotions regarding the highly human-like in more “advanced” tasks (such as lecture students at school) could be attributed to the belief that human-looking robots are expected to possess greater human competence (Jung et al., 2021).

Another trend regarding the emotional responses for all social robots, concerns how individuals express negative emotions for the healthcare condition compared to any other area. As mentioned earlier, a possible explanation is the perceived sensitivity and importance of the healthcare area. Especially since the task involves surgery, which involves potential risks and high stakes making people uncomfortable imagining having their lives in the hands of a robot. The negative feelings could therefore also reflect mistrust among people towards the social robots' ability to assist a doctor during surgery. Since patient safety is crucial within healthcare, managers and developers should consider what tasks could be appropriate for a social robot to avoid negative emotions and mistrust.
5.3 Emotions Towards the Social Robots and Their Application Area Based on Barrett & Russell’s Semantic Structure of Affect

To better understand the differences in emotions towards robots and their application areas, some results from the analysis above will be illustrated and discussed based on Barrett and Russell’s (1998) semantic structure of affect. The chosen results discussed in this section are those with the highest mean difference on the pleasure dimension. The placement of each robot and application area on the scale are based on figure 12, where the vertical line illustrates the pleasure dimension (0 = unpleasant, 100 = pleasant) and the horizontal line represents the arousal dimension (0 = deactivation, 100 = activation).

Figure 12. Barrett and Russell’s (1998) semantic structure of affect. The letters x and y represent semantic components: x = pleasantness; y = activation.
5.3.1 Companionship: Otonaroid vs. Padbot

Within the context of using a social robot to take care of and accompany older adults, the highest (pleasure) mean difference was found between the most human-like robot, Otonaroid, and the least human-like robot, Padbot. As illustrated in figure 13, the pleasure and arousal means of Padbot is fairly negative, indicating a somewhat depressed feeling among the participants when using Padbot as a companion among older adults. The pleasure and arousal means of Otonaroid, compared to Padbot, is more positive and are leaning towards “contented” and “serene”. Based on this, Otonaroid could be considered a more favorable alternative. However, the reported emotions towards it tend to be rather neutral making it somewhat difficult to interpret.

Figure 13. Padbot vs. Otonaroid in the companionship area.
5.3.2 Otonaroid: Companionship vs. Healthcare

For the most human-like robot, Otonaroid, the highest (pleasure) mean difference was found between the application area healthcare (“assist a doctor before, during and after surgery”) and companionship (“take care of and accompany older adults”). When Otonaroid was used to assist a doctor before, during and after surgery, the participants expressed more unpleasant feelings, compared to when the robot was used to take care of and accompany older adults.

Based on the model, the results indicate a tendency towards feelings of sadness or depression in the context of healthcare (figure 14). The companionship area appears to elicit more pleasant feelings, leaning towards a “contented” and “serene” feeling. However, it is important to note that the reported feelings remain relatively neutral within the context of companionship.

If a human-like social robot, such as Otonaroid, were to be used in any of the application areas, it would be less preferable and elicit more negative feelings in the healthcare context.
5.3.3 Padbot: commerce vs. Companionship

For the least human-like robot, Padbot, the highest (pleasure) mean difference was found between the application area companionship (“take care of and accompany older adults”) and commerce (“assist with navigation and payment when shopping”). In the realm of commerce, the participants expressed more positive feelings and based on the model, the means could be interpreted as a sense of relaxation or calm when the social robot was used to assist with navigation and payment when shopping (figure 15). However, even in this context, the results still leaned towards a relatively neutral positive. When Padbot was used to take care of and accompany older adults, the participants expressed significantly more negative feelings. The means for Padbot indicate a relatively unpleasant and deactivated emotional state within the companionship area, which could be characterized as “depressed” based on the model. The design of a social robot with low human-likeness as Padbot, would not be as appreciated in the area of companionship and could be more suited to be used within commerce.
5.3.4 Nao and Pepper: Commerce vs. Healthcare

The highest (pleasure) mean difference found between the application areas for both Nao and Pepper, was commerce (“assist with navigation and payment when shopping”) and healthcare (“assist a doctor before, during and after surgery”). The means of pleasure and arousal for commerce indicates a neutral but somewhat relaxed feeling, when Pepper and Nao is used as assistance within commerce. The results related to healthcare, however, suggest a greater experience of unpleasant feelings based on the model, such as sadness or depression (figure 16). When interpreting the result, it seems that Nao and Pepper would not evoke positive feelings within healthcare, thereby reducing their chances of being accepted. Just like Padbot; Pepper and Nao evoke more positive feelings within the realm of commerce. It is possible that people may be more content with the presence of these robots in that context.

5.4 Method

In the current study, a quantitative survey study was employed and published on Prolific in order to collect responses. There are several advantages of using a quantitative method to examine people's experiences of social robots. Firstly, it enables reaching a broad group of respondents to obtain a representative result. Using surveys is also a cost-effective way to gather information in research. Furthermore, a quantitative analysis method in this context provides the opportunity to identify statistically significant patterns and relationships. These patterns and relationships could hopefully be considered when developing social robots or making decisions regarding their use, since the results represent a reasonably large sample size. The current sample was collected via Prolific, based on a set of requirements; fluent in English, located in the United States, a minimum approval rate of 95% and a minimum of 100 previous submissions. These conditions should increase the reliability of the responses (Peer et al., 2022). Considering the number of participants and their reliability, it is expected that the results would be similar if the study were replicated. However, as the study did not account
for cultural differences and is geographically homogeneous, the findings should be cautiously generalized to the rest of the world until replications are conducted in other countries as well.

The study aimed to investigate people's experiences with social robots with varying degrees of human-likeness across different application areas. Using a survey as the research method relies on the participants' ability to accurately imagine the different situations and provide honest answers about their feelings. The approach carries the risk that the survey may not fairly capture the complexity of real-life situations, thereby making it challenging for participants to empathize with the situations and respond in a way that accurately reflects their feelings if they were actually exposed to those situations in real life. It would perhaps be beneficial to investigate these aspects in real-life situations. However, that requires access to social robots with varying degrees of human likenesses and the ability to simulate the different scenarios. Unfortunately, due to practical constraints and the scope of the study, conducting another method was not feasible. Although, the results indicate how the participants would hypothetically feel in the specific context.

An important limitation to consider in the present study are the given labels of each application area. While the task descriptions aim to capture significant aspects of these areas, they may not fully represent their entirety. For instance, within this study, the label “commerce” refers to the provided task description: “assist with navigation and payment when shopping”. Therefore, it is crucial to acknowledge that the study’s results cannot be generalized to all tasks within healthcare, education, commerce and companionship. The participants might have experienced different emotions if the conditions were formulated as more or less “extreme”. For instance, the healthcare condition was presented as “assist a doctor before, during and after surgery”. As mentioned in the discussion, social robots within healthcare could also be utilized for tasks such as “making beds and serving food”. This alternate formulation could potentially create more positive reactions among the participants. So when interpreting and drawing conclusions regarding the results, consider how the results are directly linked to and influenced by the formulation of the application area.

The displayed images of the social robots aimed to illustrate different degrees of human-likeness to investigate how it affects people’s emotions in the application areas. Therefore, the chosen pictures were carefully picked from a database consisting of social robots with human-likeness ranked from 0-100 (The ABOT Database, n.d.). The least and most human-like robot from the database scored reasonably similar results as in the database. However, Pepper and Nao scored less. This could be because the database ranked the human-likeness based on body-manipulators, face, and surface. Since the participants viewed all the images of the social robots simultaneously, it is possible that they compared the robots with each other. As one of the robots closely resembled a human, participants may have rated Pepper and Nao as having lower human-likeness than expected because they compared it with the highly human-like one. Since Pepper and Nao did not score a mean around 50, it carries implications for drawing conclusions regarding medium-high human-like social robots in the present study. However, the findings for the two robots are still relevant as they are well known in the research field (e.g. Thunberg et al., 2017; Manzi et al., 2021; Mubin et al.,
2018). When examining the impact of social robots' appearances on people's emotions, it is important to consider how any biases may affect the result. For instance, Tay et al. (2014) showed that individuals were more accepting of social robots that conformed to their stereotypes. In the current study, the participants reported more positive feelings towards Pepper and Nao in the commerce context compared to the healthcare context. This may be due to pre-existing stereotypes that those robots, in terms of their appearance, are better suited for commerce, rather than due to their level of human-likeness.

The current study used emotions, more specifically; pleasure and arousal based on the Affective Slider (AS), to predict how accepting people will be towards the social robot and its application area. The reason for measuring pleasure and arousal based on the AS, was motivated by its simplicity and intuitive nature. Additionally, in terms of reliability, the AS are a validated self-reporting tool for measuring human emotions (Betella & Verschure, 2016). Understanding human emotions is crucial in predicting acceptance of specific stimuli, for instance the willingness to interact with the robot (Nomura et al., 2008). Research on human emotions related to social robots are crucial to ensure that people find the development and implementation of technology pleasant. However, it would be beneficial to replicate this study and include additional questions about the participants “intention to use” the robots to further support the results.

5.5 Future Studies

The current study has contributed with insights concerning what effect the human-likeness and application area of social robots have on people’s emotions. To address what emotions certain robots and their area of use evoke, can contribute to fostering more positive attitudes regarding the development of the technology in society. The results in this study can contribute with guidance for robotic engineers, designers or policymakers since the study contains how the appearance of the robot affects emotions and which areas of use should possibly be avoided for certain robots to avoid negative emotions. The results demonstrate the importance of considering how the appearance and application area of the social robot influence social acceptance.

In future studies, to gain additional insight into social acceptance, a replication could be done where the researcher incorporates questions about the participants’ intention to use the robot in specific application areas. The current study provides information about how people may experience social robots across different situations. Consequently, the results suggest that researchers should further explore how individuals respond to robots in real-life interactions within these application areas, including what possible human behaviour it evokes and the potential societal or individual consequences it may arise.

This study did not examine the potential ethical concerns associated with social robots and their application areas. The findings suggest that people prefer a highly human-like social robot within the companionship domain. However, it is essential to consider the ethical consequences that may arise, such as the risk of people believing that human company could
be replaced with a humanoid robot. Future studies should address these ethical concerns to prevent problematic and overly positive attitudes towards the use of social robots. By exploring ethical aspects, researchers can contribute to safe and responsible integration of social robots.
6. Conclusions

In summary, the study has explored the impact of social robots' appearance, human-likeness and application area on people’s emotions, more specifically: pleasure and arousal. The aim of the study was to provide a better understanding for how the appearance, human-likeness and application area of social robots impact human emotions (and thereby acceptance). The results indicated that the social robot and the application area affected human emotions, on both the pleasure and arousal dimension. When the social robot is used to “take care of and accompany older adults” (companionship), people tend to express more positive emotions towards social robots with higher human-likeness (Otonaroid and Pepper). People seem to feel depressed when a social robot with low human-likeness (Padbot) is used within companionship. Highly human-like social robots, such as Otonaroid, elicited more positive emotional responses within the context of lecturing at school and accompanying older adults (compared to assisting a doctor). Positive emotions towards highly human-like robots in educational and companionship settings may arise from the belief in their greater human competence due to their human-like appearance. Less human-like social robots, such as Nao, Pepper, and Padbot, were associated with greater positive emotions (calm and relaxation) when used in the commerce-related task: "assisting with navigation and payment when shopping". It suggests that a more simplistic appearance with lower human-likeness may be more suitable and positively perceived for commerce-related tasks. Negative emotional responses were expressed for all robots within the healthcare condition, where the social robot was used for surgery assistance. This may reflect mistrust in the ability of social robots to assist during critical procedures, and the perceived sensitivity and importance for the particular task. The findings emphasize the importance of considering appearance and the specific tasks in designing and implementing social robots for successful integration and acceptance among humans.
References


Prolific · About Us. (n.d.). Retrieved March 30, 2023, from [https://www.prolific.co/about](https://www.prolific.co/about)


Tay, B., Younbo, J., & Taezoon, P. (2014). When stereotypes meet robots: The double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior, 38*, 75-84. [https://doi.org/10.1016/j.chb.2014.05.014](https://doi.org/10.1016/j.chb.2014.05.014)
