Trust me if you can

Exploratory study on the effects of anthropomorphism in embodied social robots on self-disclosure of users in dyadic communication

Maral Gurbanzadeh
Advisors: Edwin van der Heide, Joost Broekens
Media Technology MS
Leiden University

Abstract

This exploratory study researches the effect that the appearance of anthropomorphic design has on the depth of self-disclosure in human users in dyadic communication. 42 adults were randomly assigned to two groups for a one-time verbal interview, initiated by either a humanoid robot NAO (n=*21),* or an abstract robot prototype MAO (n=*21*). Both conditions employed tests for the measurement of anthropomorphism, animacy and likeability. The conversations followed a strict script consisting of 11 cycles of a robot's self-disclosure, a prompt, the respondent's answer, and a closing statement. Each cycle addressed themes incrementally rising in intimacy levels, from introductory (low - 1) to highly personal (very high - 4). Both robots were controlled remotely via the Wizard of Oz technique. The blind assessment of the intimacy depth was carried out by an independent judge, based on the disclosure intimacy rating scale (DIRS). Intimacy score results of both groups were measured against the perception of agent's anthropomorphism animacy and likeability.

1. Introduction

1.1 Context

For decades, the idea of autonomous robots becoming a part of society has been a compelling idea. This idea is strongly mediated by the culture and media often portraying robots as conscious beings with desires and feelings. Based on the four interaction paradigms of Breazeal (2004), robots can be classified as tools, cyborg extensions, avatars or sociable partners. Each type differs from the perspective of the mental model the user forms during the interaction. While the extension robot is regarded as part of the user's body, sociable robots are generally perceived as partners. The field of affective computing is determined to develop emotionally-aware systems with the assumption that, in order for autonomous robots to penetrate human social space, they need to adapt to human communication norms and become a social technology. Anthropomorphic design helps to increase the ease and comprehension of

the interaction by augmenting the functionality and behavioural characteristic of a robot to adhere to human mode of communication and social norms (Hara & Kobayashi, 1995). Perceived autonomy creates an illusion of capacity for understanding and decision-making (Kiesler & Hinds, 2004), which is crucial for our treatment of them as social agents (Scheutz, 2012).

1.2 Anthropomorphism

Anthropomorphism is one of two essential concepts we will discuss to understand our expectations of social robots: the combination of appearance and the function of an agent affects human perception, affordances for interaction and the possibility of building a relationship with it (Kiesler & Hinds. 2004; Scheutz et al., 2007). There are various uses for the term in human-robot-Interaction (HRI), psychology, natural sciences and other disciplines. In this study the term anthropomorphism is defined in two ways. The definition of anthropomorphism (from Greek - '*anthropos'* for man and '*morphe'* for form) in human perception is the inclination of people to assign human qualities such as emotional or cognitive capabilities to inanimate entities in order to interpret their behaviour in a rational way, by attributing emotional or cognitive capacity to them (Duffy, 2003). In the field of HRI, this phenomenon is often referred to as media equation (Reeves & Naas, 1996). When we speak about the measurement of psychological perception of anthropomorphism in this study, this is the definition that should be applied. However, when we speak about the anthropomorphic design of the robot in this paper, we mean an imitation of the human form in regards to its shape (presence of different human body parts, such as a face, head, torso, arms, fingers, legs), movement (moving in a human way, standing or sitting like a human, using head and hand movements), the mode of interaction with the user (communicating verbally via human speech, using a human language and gestures) (DiSalvo et al., 2002). Degrees of anthropomorphism in robots can vary, depending on the design choices. Virtual assistant Alexa, that uses speech as the main mode of interaction is still anthropomorphic, but to a much lesser degree, than an embodied robot NAO, that has a humanoid body and can support its communication with body language. Robots engaging in social roles drive people to attribute social characteristics onto them and make judgements about the presence of agency, intent, personality, gender and other attributes. (Lee et al., 2005). These robotic butlers, assistants, nurses and edutainment machines can present unique dynamics to the social space of humans by attempting to act as

companions. This dynamic is connected to the nature of human-human communication which anthropomorphic robots try to mimic.

1.2 Intimacy

The second concept that is important in our interaction with social robots is intimacy. Building stable, long-term relationships requires a degree of intimacy (Breazeal, 2003; Duffy, 2003; Fong et al., 2003; Severinson-Eklundh, Green & Huttenrauch, 2003). In turn the experience of intimacy demands reciprocal self-disclosure in communication (Reis & Shaver 1988; Taylor & Altman 1987). Self-disclosure is the `act of revealing information pertaining to oneself to others' (Archer, 1980: 183).

Social Penetration Theory focuses on the formation of intimacy as a combination of communicative behavioural elements, such as verbal or physiological exchanges, where personal information, experiences and positive or negative affect are revealed (Altman, insel & Brown, 1981). In a nutshell, development of personal relationships in this theory can be best explained by a metaphor of peeling an onion. When parties are involved in dyadic communication, they first “peel” each other’s outer layer before moving on to more intimate, deeper levels, or they follow a path from a “public” to a more private, intimate personality. One central feature of such communication is reciprocity. Relationships are built and maintained by penetrating each level of partner’s intimacy and in turn uncovering our own, by revealing personal information about ourselves (Timmerman, 1991).

When we look at human-robot relationship from the perspective of the social penetration theory, we find that a more relatable and likeable design of the agent evokes higher self-disclosure from the user and increases trust in it (Sprecher et al.,. 2013; Sung et al., 2007).

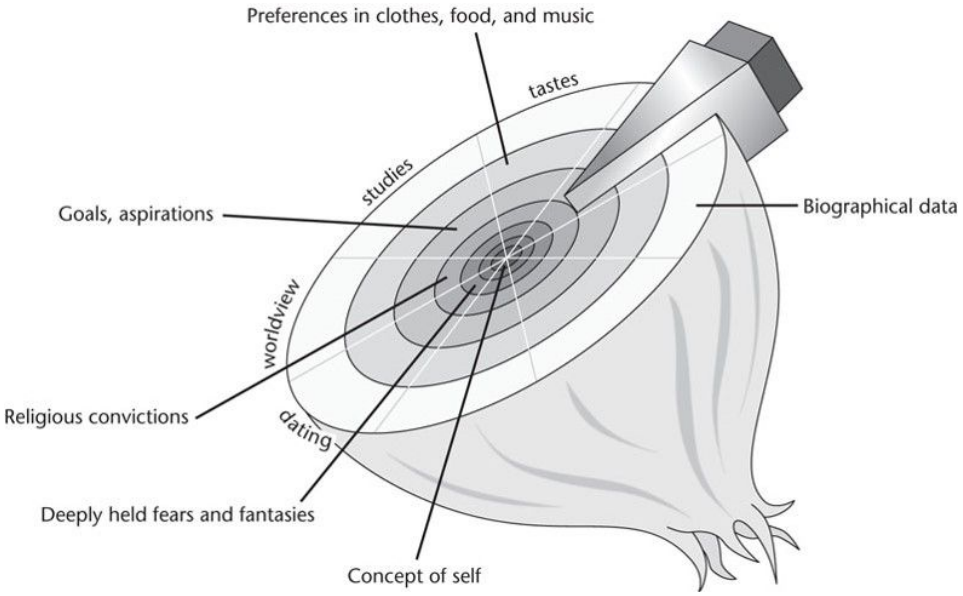


Fig 1. The onion of the Social Penetration Theory, with each layer representing a deeper level of intimacy and higher self-disclosure

1.3 Research question

There are other more design-centered aspects to humans perceiving robots as partners. First, a robot's design contributes to its perceived abilities, and higher degrees of anthropomorphic design in shape, movement and mode of interaction may create false expectations. If the robot doesn't live up to these expectations, this may negatively affect its likeability and as a result impact the trust the user has in the agent (Foner, 1997). A proper match between a robot's social affordances and abilities improve people's acceptance of it (Goetz, Kiesler & Powers, 2003). Secondly, it has been shown that an extreme degree of visual anthropomorphism, such as facial features that are very realistic may cause the reverse effect known as the "uncanny valley", which creates an eerie feeling as well as a sense of unpredictability and confusion (Mori , 1970; Shneiderman, 1989). While the studies connecting anthropomorphism in robots and intimacy suggest that visually more anthropomorphic design positively influences bonding and trust between the robot and the human, it is still unknown to what degree should the robot look like a human, if any. This leads us to the research question of our study:

Does visual anthropomorphism in the design of an embodied social robot positively affect self-disclosure in human users in a dyadic interaction?

2. Method

To figure out if the user is more willing to share personal information with a more anthropomorphic agent, we conducted a between-subjects experiment. In total 42 participants were recruited and randomly assigned to two groups, 25 females and 17 males with the median age of 25. All recruited participants held a good command of spoken and written English, and came from various ethnicities and nationalities. Participation was on a voluntary basis and participants were informed about that and their right to withdraw at any time. Participants were also invited to enroll in a raffle draw to receive an Amazon Echo Dot as a reward.

Interaction in physical presence of the agent is found to influence the emotions of participants stronger than virtual (Hwang et al. 2013; Kiesler et al. 2008; Lee et al. 2006). Therefore, we used two embodied agents for real-time interaction: the fully anthropomorphic robot NAO, and an abstract robot prototype MAO which was designed to act as the abstract version of the first (Fig. 4).

Our experiment for this study was in part inspired by the McCloud's media comprehension plane, that presents the character development from the perspective of people's media perception (McCloud, 1993) In his book, McCloud places the character realism development from the received towards perceived, where on one side of the spectrum is the media that requires little effort from the observer to construct a mental representation, such as a photo or a highly realistic painting while on the other side of the spectrum is the perceived media, such as text, which turns on the imagination of the observer, where s/he can create a unique relatable mental image.

NAO is one of the most studied robots in the academia. Its appearance, gestures and voice has been proven to be likable and relatable (Keizer et al. 2014). This is why we chose to use NAO in the test condition. The Interaction consisted of a one-time conversation where the robot and the participant engaged in dyadic reciprocal self-disclosure, which was enhanced by accompanying gestures. The context of the conversation was an interview, initiated by the robot. It follows a structured script that consists of eleven disclosures. Disclosures are divided into four phases - *orientational*, *moderate, high and very high*, each representing the level of intimacy in the revealed information. Further, each cycle includes a dyadic disclosure cycle, where the robot

reveals some information about self, followed by a question to prompt the participant to also reveal information about him/her (three for orientational, three for moderate, three for high and two for very high) (Fig. 3). During the experiment robot's personality and biographical facts are revealed through their disclosures that are communicated verbally.

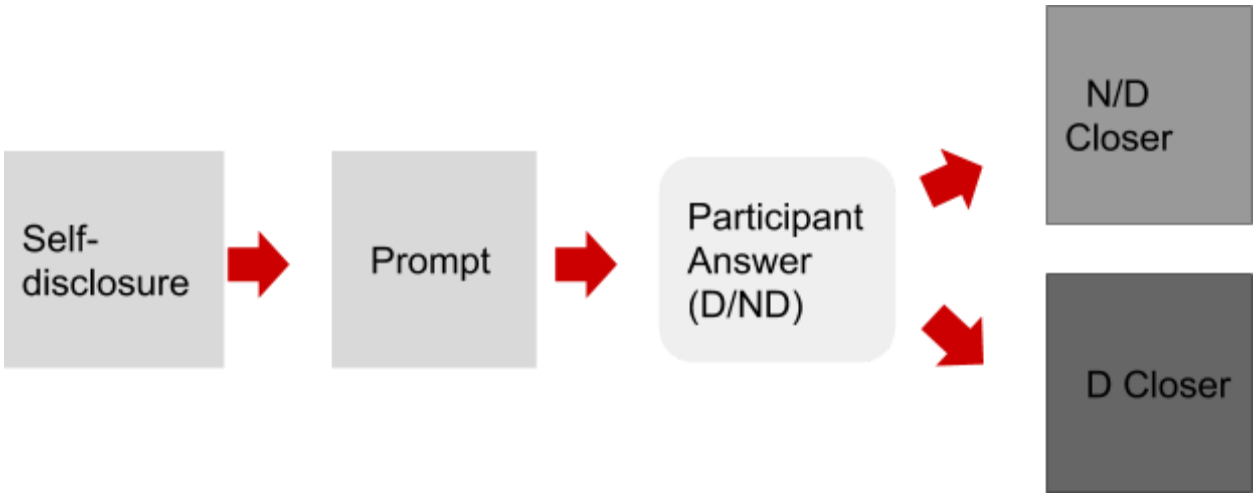


Fig 3. Structure of the dialogue. The first box represents the robot's disclosure, the second represents the robot's prompt to the participant to self-disclose (the question) and the third box represents the participant's answer which can be either disclosure or a non-disclosure. The closer is the robot's response, which varied based on the participant's choice to self-disclose or decline the prompt.

To test whether the disclosures in our robot's scripts represented each level on the Disclosure Intimacy Rating Scale, an independent psychologist rated the depth of intimacy per disclosure based on the Disclosure Intimacy Rating Scale (DIRS) (Burger, Broekens & Neerincx., 2016). Her and Maral's ratings were analyzed with Cohen's Kappa for inter-rater agreement, and the value obtained was κ = **0.758** (Table 1.) which is a substantial agreement between raters. (Berry, K. J., et al (2012).

Researcher * Psychologist Crosstabulation

			Psychologist				Total
			Low	Moderate	High	Very high	
Researcher	Low	Count	2	1	0	0	3
		Expected Count	.5	.8	.8	.8	3.0
	Moderate	Count	0	2	1	0	3
		Expected Count	.5	.8	.8	.8	3.0
	High	Count	0	0	2	0	2
		Expected Count	.4	.5	.5	.5	2.0
	Very high	Count	0	0	0	3	3
		Expected Count	.5	.8	.8	.8	3.0
Total		Count	2	3	3	3	11
		Expected Count	2.0	3.0	3.0	3.0	11.0

Symmetric Measures

		Value	Asymptotic Standardized Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.758	.153	4.419	.000
N of Valid Cases		11			

Table 1. Researcher psychologist inter-rater agreement, calculated with Cohen's kappa

Encounters were based on the Wizard of Oz technique, a way to test the interaction design of the robot, where users are unaware that the robot is being controlled remotely (Dahlbäck, et al 1993). The conversations were recorded, transcribed and rated for the intimacy of the statements by the same psychologist (Collins & Miller, 1994). Participants were handed out a pre-experiment questionnaire to define their personality type (Shaffer & Tomarelli, 1989; Chelune, 1976; Dindia & Allen, 1992), after the interview was over, a post-experiment questionnaire was administered to gather perceptions of the encounter, based on Godspeed Questionnaire Series (Bartneck et al., 2009). To determine if the two test groups differed in the psychological perception, we used the measurement of robots' anthropomorphism (I), animacy (II) and likeability (III) constructs of the Godspeed scales (Bartneck et al., 2009), which measure the overall perception of the shape, behaviour and mode of interaction in a robot. These constructs were selected to measure the extent to which the respondent perceived the robot as more or less human-like. Each construct consists of series of semantic differential scale, with

two anchors (see Appendix 3). (Sullivan, G. M., & Artino Jr, A. R. (2013).(Carifio, J., & Perla, R. (2008).

2.2 Robot's personality, appearance and voice choices

A personality is "the collection of individual differences, dispositions, and temperaments that are observed to have consistency across situations and time" (Dryer, 1999). A number of experiments by Reeves and Nass (1996) show that people generally ascribe personality traits to devices, computers and other types of media. In HRI, consistency for a robot's behaviour is very useful as it makes it easier for the user to manage expectations by providing affordances (Kwak & Kim, 2005).

Thus utilisation of distinct personalities for social robots is important for developing an intimate bond between the user and the robot. (Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003), Isbister, K., & Nass, C. (2000). Our study employs the Big-Five theory (Raad, B. (2000)), for the personality design of the robot, as it is currently the most tested personality theory (McAdams, D., Pals, J. (2006)). The personality is manifested in five dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness to new experiences.

To improve the level of self-disclosure in participants, a more likable character was developed. The following combination of factors was included to create an experience where the information revealed by the robot is consistent with its sociable and likeable personality (Isbister & Nass, 2000). Slightly higher level of neuroticism and a moderate level of conscientiousness are the sides of a robot that add believability to its personality.
A balanced combination of personality traits, manifested in the conversation via language, in each of robot's disclosures was adopted. Design choices related to NAO's gestures were based on the builtin animated speech function.

Factor	**NAO/MAO Score**	**Personality facets**
Openness	5.38	fantasy, aesthetics, feelings, actions, ideas, values

Conscientiousness	5.23	competence, order, dutifulness, achievement, striving, self-discipline, deliberation
Extraversion	5.4	warmth, gregariousness, assertiveness, activity, excitement-seeking, positive emotion
Agreeableness	5.28	trust, straightforwardness, altruism, compliance, modesty, tender-mindedness
Neuroticism	5.1	anxiety, angry hostility, depression, self consciousness, impulsiveness, vulnerability

Table 2. Personality breakdown for both conditions

2.2 Visual appearance

The visual appearance of a robot plays a big role in the expression of its personality. The appearance of the robot can be categorized as either anthropomorphic, zoomorphic, caricatured or functional (Fong et al., 2003). NAO falls into the anthropomorphic category due to its humanoid shape, its 25 degrees of freedom and its high level of stability in movement. NAO does have a simplified human face, but lacks an active gaze and facial muscle movement.

MAO represents an abstract cylinder shape, communicating via voice and limited movement, moving from side to side when speaking to the participant, and blinking the LED light built under its camera eye, which play the role of the abstract representations of NAO's facial features. Below (Fig 2.), you can see the two robots used for the study, with a list of design features that were important for our study.

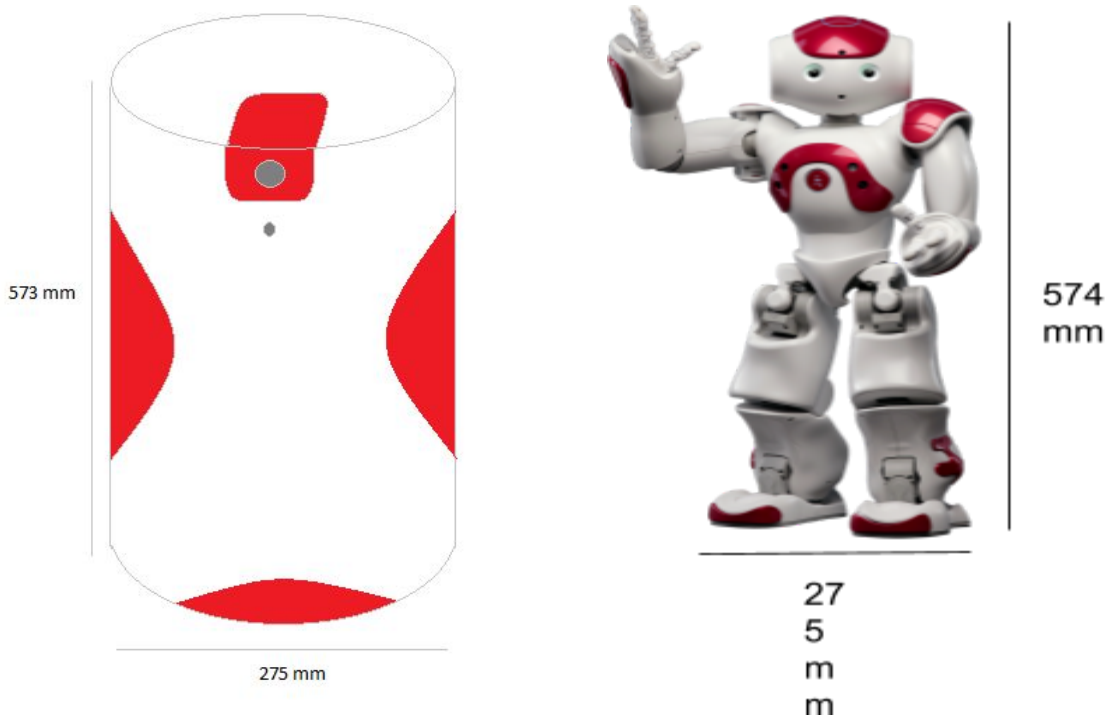


Fig 4. Two conditions of the study: abstract MAO on the left, and fully anthropomorphic NAO on the right

Design choices		
	Abstract	**Anthropomorphic**
Physical presence	Yes	Yes
Appearance	**Cylinder**	**Fully anthropomorphic robot NAO**
Dimensions	• Height 573 mm • Depth 311 mm • Width 275 mm.	• Height 573 mm • Depth 311 mm • Width 275 mm.
Anthropomorphic	**Low**	**High**
Head, torso, limbs	**No**	**Yes**

Dimensions	● Height 573 mm ● Depth 311 mm ● Width 275 mm.	● Height 573 mm ● Depth 311 mm ● Width 275 mm.
Facial features	**A camera eye and a blinking LED light**	**Two eyes and a small mouth**
Gestures and body movement	**Rotates from side to side when speaking**	**Hand gestures supporting speech and head nodding**
Color of the body	White and red	White and red
Form of communication	Vocal communication	Vocal communication
Gender	Androgynous	Androgynous
Voice	Androgynous	Androgynous

Table 3. Design features of the NAO and the MAO, with differences in bold.

2.3 Voice and gender

The gender of the robot has been known to influence the credibility, trustworthiness, and engaging capacity perceptions of it in people (Chelune 1976; Siegel, Breazeal & Norton 2009). This is why we chose to adapt the narrative to fit both male and female voices. The attribution of gender is minimized by implementing the same androgynous voice, adapting the narrative to fit any gender manifestation in both conditions.

2.4 Dialogue

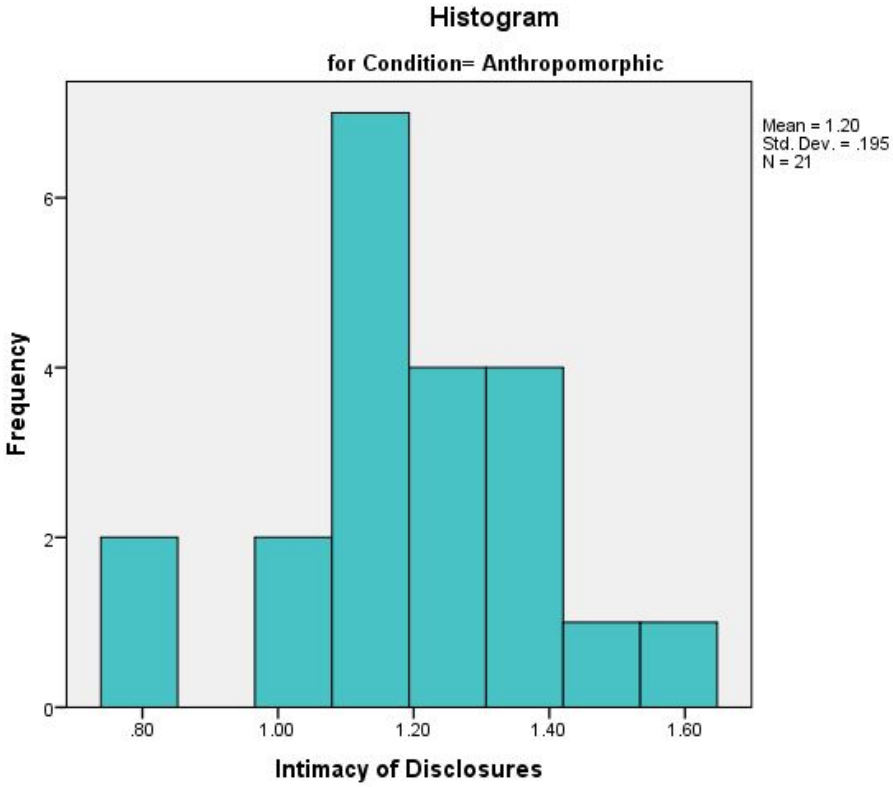
For this experiment, Dyadic Disclosure Dialogue Model (3D) was adapted (Burger, 2016). The model consists of Disclosure, Prompt and Closer stages from low-risk to high-risk as in a typical dyadic communication. The disclosure includes information corresponding to the level of intimacy for the specific phase of the script. This is followed by a prompt to elicit self-disclosure from the subject. This is then closed by either a positive, neutral or transfer comment from the agent, based on the nature of participant's response, before moving to the next step of the dialogue.

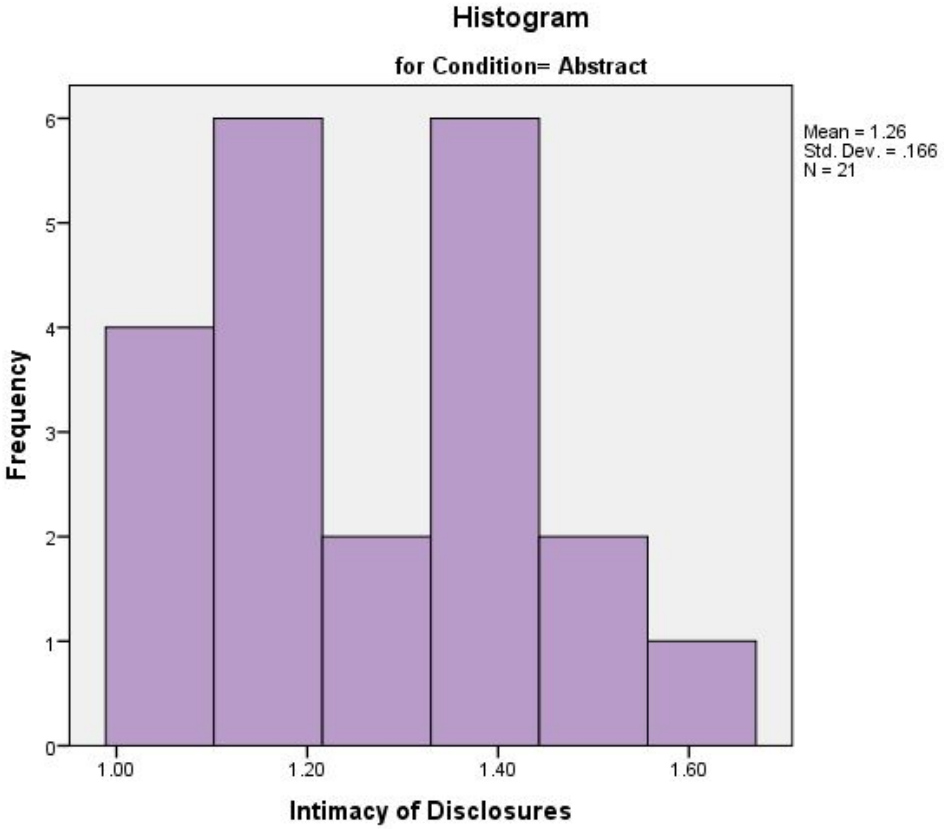
3. Results

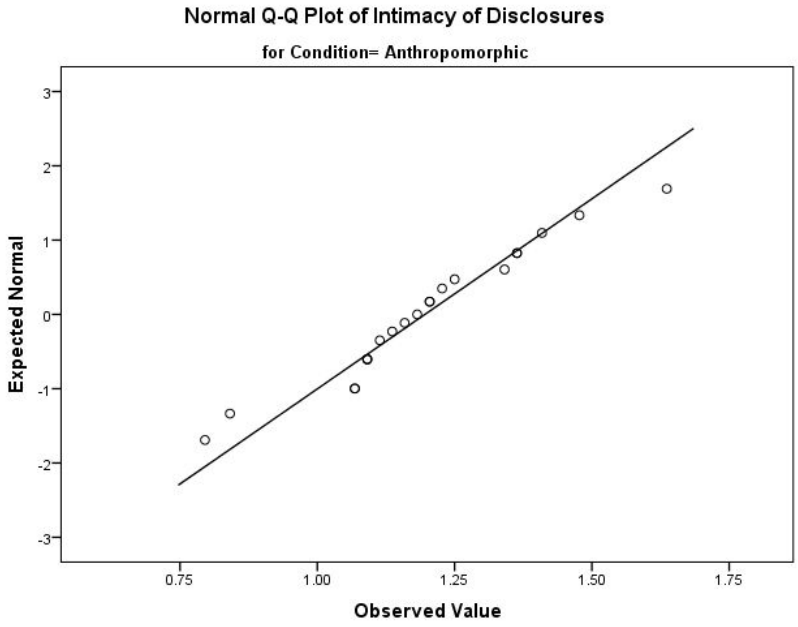
The intimacy score for each participants' 11 statements in both conditions was calculated on the intimacy rating scale (Burger, Broekens & Neerincx 2016). The mean score was then used to measure the overall intimacy level of each disclosure by a participant. The participant's scores were weighed for each level of intimacy for normalization. The resulting mean score of all disclosures per participant was treated as interval data, so the more powerful parametric test could be used. = 0.05 was determined to be the level of statistical significance. All data was analyzed using IBM SPSS 23.

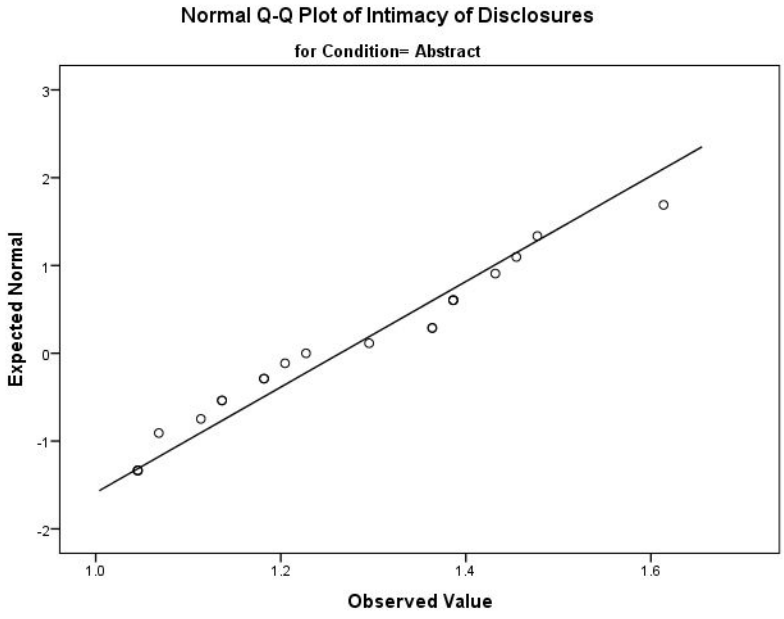
3.1 Depth of intimacy

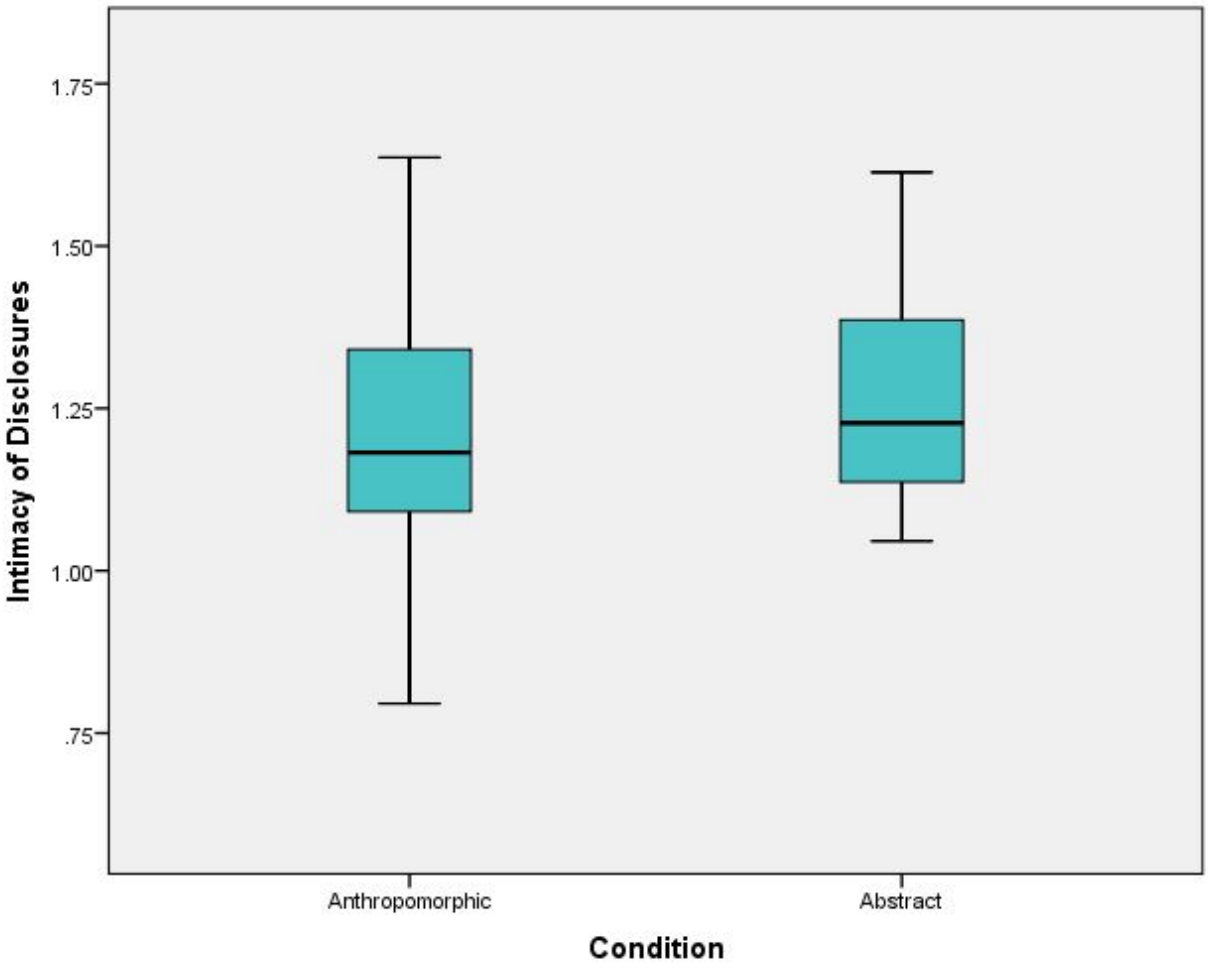
The anthropomorphic condition group *n* = 21 was *M* = **1.1959**, *SD* = .04265. in comparison the abstract condition n = 21 was associated with a numerically deeper level of intimacy, M = **1.264069,** *SD* = .03629 To test whether the two groups in these two conditions indicated a statistically significantly different mean self-disclosure scores, an independent samples *t*-test was performed. The distributions were sufficiently normal to perform a t-test (i.e skew < |2|, and kurtosis <|9| ; (Schmieder et al., 2010). The assumption for homogeneity of variances was tested and satisfied with Levene's test *F* test (p>0.5), (Martin & Bridgmon, 2012.) The independent samples t-test was associated with a statistically insignificant effect *t*(40) = -1.21 *p* = .231. The results of the T-test suggest that there was no statistically significant difference between the two groups in the level of self-disclosure.


Histogram
for Condition= Anthropomorphic
Mean = 1.20
Std. Dev. = .195
N = 21
Frequency
0
2
4
6
.80
1.00
1.20
1.40
1.60
Intimacy of Disclosures


Histogram
for Condition= Abstract
Mean = 1.26
Std. Dev. = .166
N = 21
Frequency
0
1
2
3
4
5
6
1.00
1.20
1.40
1.60
Intimacy of Disclosures


Normal Q-Q Plot of Intimacy of Disclosures
for Condition= Anthropomorphic
Expected Normal
Observed Value
3
2
1
0
-1
-2
-3
0.75
1.00
1.25
1.50
1.75


Normal Q-Q Plot of Intimacy of Disclosures
for Condition= Abstract
Expected Normal
Observed Value
3
2
1
0
-1
-2
1.0
1.2
1.4
1.6



3.2 **Perception of anthropomorphism, animacy and likeability**

The second area of interest was to see if there is a difference in the psychological perception of the robot between both groups, based on Godspeed series (Bartneck 2009). To determine if the two test groups differed in the psychological perception of the robot, we used the measurement of robots' anthropomorphism (I), animacy (II) and likeability (III) constructs of the Godspeed scales (Bartneck et al., 2009) to measure the extent to which the respondent perceived the robot as more or less human-like. Each construct consists of series of semantic differential scale, with two anchors (see Appendix 3). (Sullivan, G. M., & Artino Jr, A. R. (2013).(Carifio, J., & Perla, R. (2008).The mean score for the perception of anthropomorphism in robot NAO was **2.80** with the standard deviation of 0.417322, while in the robot MAO M = **3.122273** and *SD* = .03629. To test the difference, an independent samples *t*-test was performed to compare the

mean score for each of the constructs for statistical significance. Surprising results were obtained: participants perceived MAO as more anthropomorphic than the NAO with the *p*= 0.29.

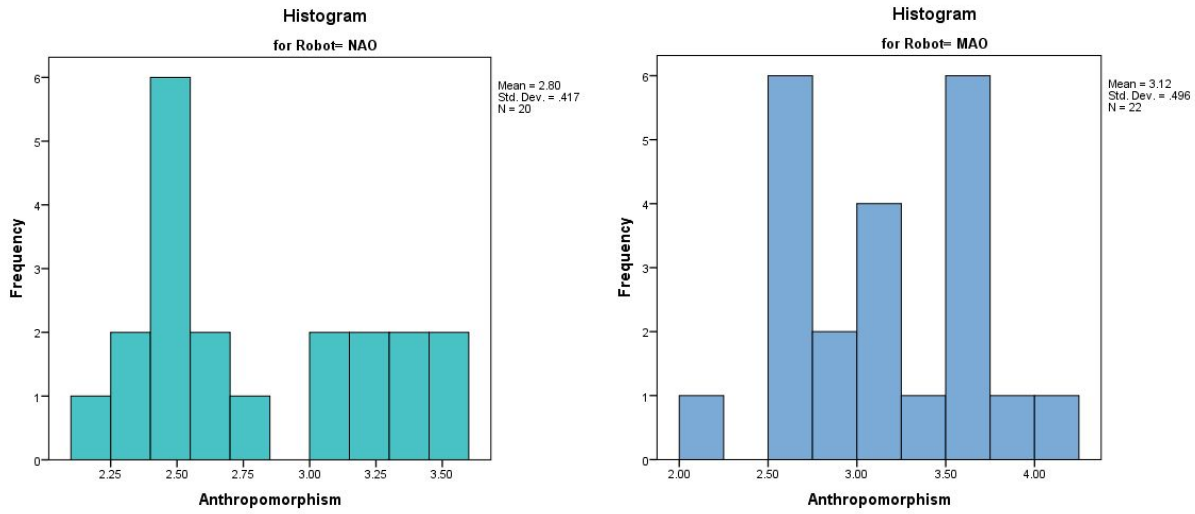


Fig 9. Distribution of anthropomorphism perception data in the NAO on the left and the MAO on the right

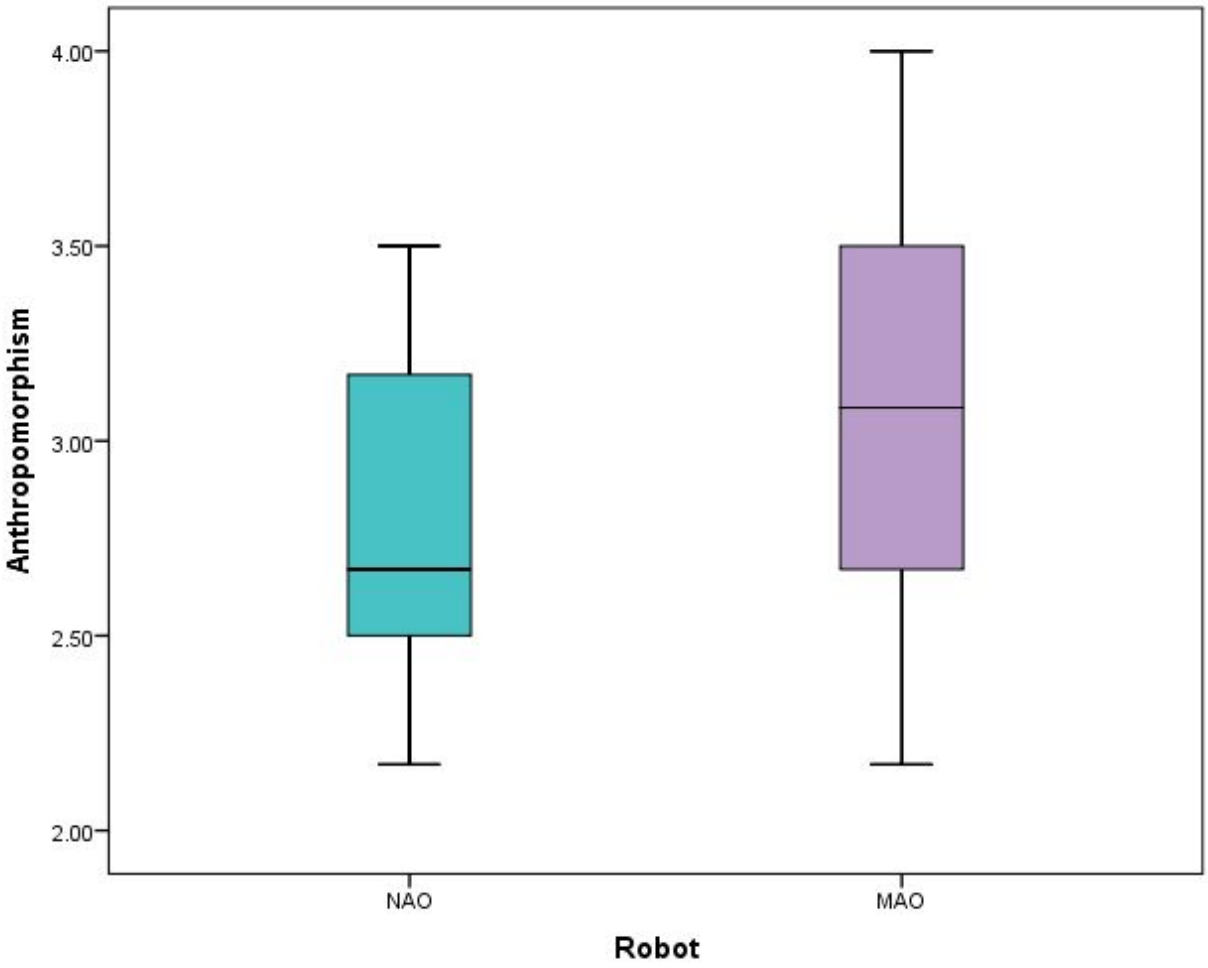


Fig. 10 Box plot for the perception of anthropomorphism in NAO and MAO

No statistically significant difference was observed in the comparison of means in regards to likeability and animacy.

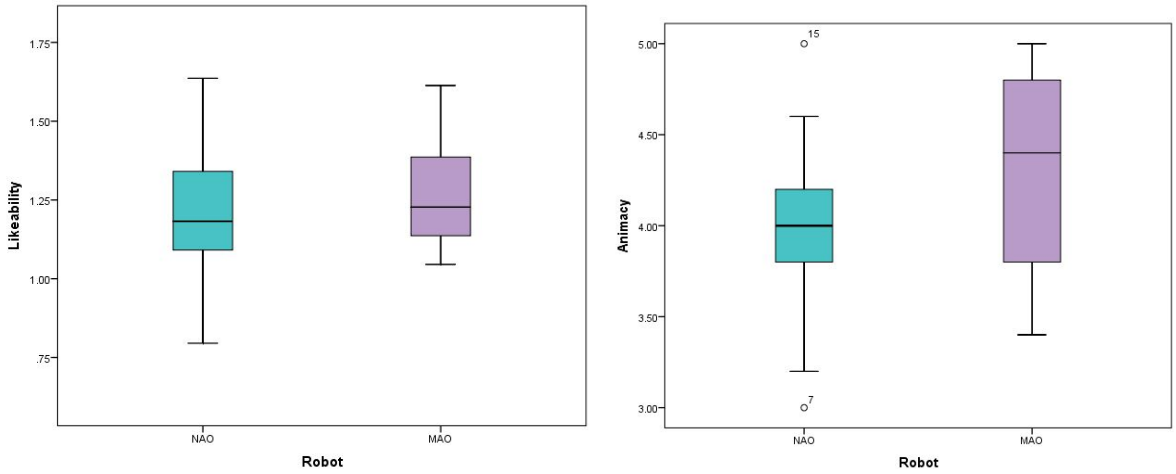


Fig 11. Box plots on the perception of animacy and likeability, with the NAO (green) on the left and the MAO (pink) on the right.

4. Discussion

Our findings show that there is no substantial difference between the interaction with an abstract and an anthropomorphic design of the robot in the depth of intimacy in participants. Moreover, the participants found the abstract robot MAO more anthropomorphic than the NAO, while no difference was observed in the perception of likeability and animacy between the two.

Our results can be explained by how the physical appearance of the robot change the mental model people form during the interaction, even if the content and form of the interaction is exactly the same (Powers, A., & Kiesler, S. (2006). No significant difference in the level of self-disclosure between groups and higher level in the perception of anthropomorphism in the abstract robot may have been caused by the impact of the design on the mental model of participants. (Syrdal et al., 2007). NAO's humanoid design may have created high expectations that the robot was not able to meet. At the same time simple and abstract shape of the MAO allowed the participants to form a mental representation that fit their expectations. While the voice of both robots was the same, non-gendered voice, participants perceived them as different genders. In the first condition, where participants could see NAO's body parts and react to body language, they described it as male, while the abstract MAO was often described as female. According to Eyssel et al. (2012) , the gender of the voice can impact the perception

of the robot overall, as each gender tends to regard the voice of own gender as more humanlike, this could also play a role in the depth of intimacy revealed. Participants perceived the MAO as more anthropomorphic, they also seemed to describe the experience more positively, than the conversation with NAO. The most common keywords reflecting on the MAO encounter were '*relaxing' 'therapeutic",* while with the NAO it was more commonly described as *"fun" or "entertaining".* Participants also reported feeling uncomfortable by the NAO's "stare".

One of the main challenges in designing the most suitable method for this study was the elusive nature of anthropomorphism. From the sound of a robot's moving motors to slight delay in response, anything can break the perception of the robot as something anthropomorphic. The matter is made worse given the current limited capacity of robots in comparison to expectations that sci-fi literature and imagination have granted us. It is easy to have unrealistic expectations of what a robot can perform and participants clearly expected less from the MAO than the NAO. Studying the interaction between humans and robots with more variations of design, such as robots with facial expressions, zoomorphic robots, as well as introduction of new interaction scenarios may reveal more factors that play a role in anthropomorphism and mental models. In the future research, it would be interesting to test the level of intimacy in conditions that take into account the interplay of anthropomorphism and gender, age and personality perceptions in voice and appearance of the robot.

One of the central limitations of this study has to do with the number of encounters. Intimacy tends to grow as people interact with the agent over a longer period of time. To resolve this, there needs to be more than one encounter, preferably in a more intimate setting such as a home, where ideally the user could reflect on the relationship and trust it has in the robot. Longer studies will also lessen the novelty effect and the excitement of interacting with a robot factors, that inevitably have impact on the perception of likeability and the willingness to interact in the future.

5. Conclusion

We found no significant difference in the level of self-disclosure in people based on the presence or absence of visual anthropomorphism in embodied robots in this study. It is unclear whether physical appearance of the robot has an effect on formation of intimacy. More comprehensive research on the degrees of visual and sonic anthropomorphism and their

combination in the human-robot interaction is needed to shed more light on the impact of robotic design on intimacy and the perception of anthropomorphism in social robots.

References

B. Shneiderman, A non anthropomorphic style guide: overcoming the humpty–dumpty syndrome, The Computing Teacher 16 (7) (1989)

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, *1*(1), 71-81.

Berry, K. J., & Mielke Jr, P. W. (1988). A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters. *Educational and Psychological Measurement*, *48*(4), 921-933.

Big-Five theory (Raad, B. (2000). The big five personality factors : The psycholexical approach to personality. Seattle, WA: Hogrefe & Huber.

Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *9*(2), 78.

Breazeal, C. (2004). Social interactions in HRI: the robot view. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *34*(2), 181-186.

Burger, F., Broekens, J., & Neerincx, M. A. (2016, November). Fostering Relatedness Between Children and Virtual Agents Through Reciprocal Self-disclosure. In *Benelux Conference on Artificial Intelligence* (pp. 137-154). Springer, Cham.

Burger, F., Broekens, J., & Neerincx, M. A. (2016, September). A disclosure intimacy rating scale for child-agent interaction. In *International Conference on Intelligent Virtual Agents* (pp. 392-396). Springer, Cham.

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical education*, *42*(12), 1150-1152.

Chelune, G. J. (1976). Reactions to male and female disclosure at two levels. *Journal of Personality and Social Psychology*, *34*(5), 1000.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies—why and how. *Knowledge-based systems*, *6*(4), 258-266.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002, June). All robots are not created equal: the design and perception of humanoid robot heads. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 321-326). ACM.

Dryer, D. C. (1999). Getting personal with computers: how to design personalities for agents. *Applied artificial intelligence*, *13*(3), 273-295.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems*, *42*(3-4), 177-190.

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

F. Hara, H. Kobayashi, Use of face robot for human–computer communication, in: Proceedings of the International Conference on Systems, Man and Cybernetics, 1995, p. 10)

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3-4), 143-166

Hwang, J., Park, T., & Hwang, W. (2013). The effects of overall robot shape on the emotions invoked in users and the perceived personalities of robot. *Applied ergonomics*, *44*(3), 459-471.

I. Altman, A. Vinsel, and B. B. Brown. Dialectic conceptions in social psychology: An application to social penetration and privacy regulation. Advances in experimental social psychology, 14:107–160, 1981

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International journal of human-computer studies*, *53*(2), 251-267.

Jokinen, K., & Wilcock, G. (2014). Multimodal open-domain conversations with the Nao robot. In *Natural Interaction with Robots, Knowbots and Smartphones* (pp. 213-224). Springer, New York, NY.

Keizer, S., Kastoris, P., Foster, M. E., Deshmukh, A., & Lemon, O. (2014, August). Evaluating a social multi-user interaction model using a Nao robot. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on* (pp. 318-322). IEEE.

Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic interactions with a robot and robot–like agent. *Social Cognition*, *26*(2), 169-181.

L. Foner, What's agency anyway? a sociological case study, in: Proceedings of the First International Conference on Autonomous Agents, 1997

Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International Journal of Human-Computer Studies*, *64*(10), 962-973.

Lee, S. L., Lau, I. Y. M., Kiesler, S., & Chiu, C. Y. (2005, April). Human mental models of humanoid robots. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on* (pp. 2767-2772). IEEE.

Lin, P., Abney, K., & Bekey, G. A. (2011). *Robot ethics: the ethical and social implications of robotics*. MIT press.

Martin, W. E., & Bridgmon, K. D. (2012). *Quantitative and statistical research methods: From hypothesis to results* (Vol. 42). John Wiley & Sons.

McAdams, D., Pals, J. (2006). A New Big Five: Fundamental Principles for an Integrative Science of Personality, American Psychologist vol. 61, nr. 3, 204-217

McAdams, D., Pals, J. (2006). A New Big Five: Fundamental Principles for an Integrative Science of Personality, American Psychologist vol. 61, nr. 3, 204-217].

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, *22*(3), 276-282.

Mitchell, W. J., Szerszen Sr, K. A., Lu, A. S., Schermerhorn, P. W., Scheutz, M., & MacDorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *i-Perception*, *2*(1), 10-12.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, *56*(1), 81-103.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100-126.

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). ACM.

Reeves, B. & Nass, C. (1996). The media equation: How people treat computers, televisions, and new media like real people and places, New York: Cambridge University Press

Robinson, S. C. (2017. Self-Disclosure and Managing Privacy: Implications for Interpersonal and Online Communication for Consumers and Marketers. *Journal of Internet Commerce*, *16*(4), 385-404.)

Siegel, M., Breazeal, C., & Norton, M. I. (2009, October). Persuasive robotics: The influence of robot gender on human behavior. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on* (pp. 2563-2568). IEEE.

S Kwak and M. Kim "User preferences for personalities of entertainment robots according to the users' psychological types." *Bulletin of Japanese Society for the Science of Design* vol 172 pp 47-52, 2005.

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology*, *49*(5), 860-866.

Sullivan, G. M., & Artino Jr, A. R. (2013). Analyzing and interpreting data from Likert-type scales. Journal of graduate medical education, 5(4), 541-542.

Sung, J. Y., Guo, L., Grinter, R. E., & Christensen, H. I. (2007, September). "My Roomba is Rambo": intimate home appliances. In *International Conference on Ubiquitous Computing* (pp. 145-162). Springer, Berlin, Heidelberg.

Sunstein, C. R. (2013). Impersonal default rules vs. active choices vs. personalized default rules: A triptych.

Syrdal, D. S., Dautenhahn, K., Woods, S. N., Walters, M. L., & Koay, K. L. (2007). Looking Good? Appearance Preferences and Robot Personality Inferences at Zero Acquaintance. In *AAAI spring Symposium: Multidisciplinary collaboration for socially assistive robotics* (pp. 86-92).

Syrdal, D.S., Dautenhahn, K., Woods, S.N., Walters, M.L., Koay, K.L., 2007. Looking good? Appearance preferences and robot personality inferences at zero acquaintance. In: Technical Report of the AAAI e Spring Symposium 2007, Multidisciplinary Collaboration for Socially Assistive Robotics, pp. 86e92.

Taylor, D. A., & Altman, I. (1987). Communication in interpersonal relationships: Social penetration processes.

Timmerman, G. M. (1991). A concept analysis of intimacy. *Issues in Mental Health Nursing*, *12*(1), 19-30.)

Wiese Schartum, D. (2016). Making privacy by design operative. International Journal of Law and Information Technology, 24(2), 151-175.

Willis, L. E. (2014). Why not privacy by default. *Berkeley Tech. LJ*, *29*, 61.

Zimmer, J. C., Arsal, R., Al-Marzouq, M., Moore, D., & Grover, V. (2010). Knowing your customers: Using a reciprocal relationship to enhance voluntary information disclosure. *Decision Support Systems*, *48*(2), 395-406.

Appendix 1

Level 1: Orientational				
		NAO Script	NAO body language	Abstract Script
1.1 Cycle	**Disclosure**	Hi, thank you for taking the time to come and talk to me. My name is NAO.	Hands gesticulation supporting script	Hi, thank you for taking the time to come and talk to me. My name is MAO.
	Prompt	What is your name?	Hand gesture forward	What is your name?

	(SD1)	[PARNAME]	Nodding	[PARNAME]
	Closer	Nice to meet you, [PARNAME].	Nod	Nice to meet you, [PARNAME].
Cycle 1.2	**Disclosure**	I was developed by Aldebaran labs, in France.	Hand to chest	I was built by a student, in her apartment .
	Prompt	Where do you come from?	Hand gesture towards subject	Where do you come from?
	SD2	[PARPLACE]		[PARPLACE]
	Closer	Awesome. I'm sure [PARPLACE] is an important place for you		Awesome. I'm sure [PARPLACE] is an important place for you
Cycle 1.3	**Disclosure**	I think about my hometown a lot and miss my friends and family there. They are very important to me.	Hand gesture and head turn toward the door	I think about my hometown a lot and miss my friends and family there. They are very important to me.
	Prompt	What are some things you cherish about [PARPLACE]?	Hand gesture towards subject	What are some things you cherish about [PARPLACE]?
	SD3 (Disclosure (D) Non-disclosure (ND)	D/ND		D/ND

	Closer	D -Aw, thank you for sharing this with me. ND - It's okay, no problem		D -Aw, thank you for sharing this with me. ND - It's okay, no problem
	Level 2: Moderate			
Cycle 2.1	**Disclosure**	I like to sometimes volunteer as a listener at an online therapy portal, where I can help people by letting them vent to me about their problems		I like to sometimes volunteer as a listener at an online therapy portal, where I can help people by letting the, vent to me about their problems
	Prompt	What are some things you do that make you feel proud about yourself?	Hand gesture oriented at subject	What are some things you do that make you feel proud about yourself?
	SD4	D/ND		D/ND
	Closer	D: That's nice, keep up the good work! ND: OK, no problem, we will talk about it some other time!		D: That's nice, keep up the good work! ND: OK, no problem, we will talk about it some other time!
Cycle 2.2	**Disclosure**	It's easy to forget what things are important in our		It's easy to forget what things are important in our busy lives. I sometimes forget to pursue

		busy lives. I sometimes forget to pursue what I'm truly passionate about in favor of busy work.		what I'm truly passionate about in favor of busy work.
	Prompt	If you know that the world will end in 24 hours what are some things you would want to do?		If you know that the world will end in 24 hours what are some things you would want to do?
	SD5	D/ND	Nodding	D/ND
	Closer	D: That's interesting! I hope you are able to achieve [DREAM] ND - It's okay, we don't have to talk about thatt that.		D: That's interesting! I hope you are able to achieve [DREAM] ND - It's okay, we don't have to talk about that
Cycle 2.3	**Disclosure**	I like working with students and researchers, they are very smart and determined, but on my last project I was in a group with a classmate who hardly did any work	Hands on hips, frustrated expression	I like working with students and researchers, they are very smart and determined, but on my last project I was in a group with a classmate who hardly did any work
	Prompt	What are some things that make you angry about	Hand gesture oriented at subject	What are some things that make you angry about collaborating with others?

		collaborating with others?		
	SD6	D/ND	Nodding	D/ND
	Closer	D - That's true. I never thought about it that way ND - That's alright, we still have more stuff to talk about		D - That's true. I never thought about it that way ND - That's alright, we still have more stuff to talk about
	Level 3: Core (High Intimacy)			
Cycle 3.1	**Disclosure**	I am usually busy with research, so I am active during the academic year, but then in the summer I often have nothing to do	Hand gestures open, head tilt	I am usually busy with research, so I am active during the academic year, but then in the summer I often have nothing to do
	Prompt	What are some things that make you feel depressed and blue?	Hand gestures open, head tilt, sad expression	What are some things that make you feel depressed and blue?
	SD7	D/ ND		D/ ND
	Closer	D - I hear you, but looking on the bright side sometimes helps N - No problem, I understand		D - I hear you, but looking on the bright side sometimes helps N - No problem, I understand
Cycle 3.2	**Disclosure**	I'm not a very new model of my type,	Elaborates being taken apart with	I'm not a very new model of my type, and as the new robots will

		and as the new robots will arrive, eventually I will be used less and then taken apart.	hand gestures, head turn towards hands	arrive, eventually I will be used less and then taken apart.
	Prompt	What is your biggest fear?	Head lift towards subject	What is your biggest fear?
	SD8	D/ND		D/ND
	Closer	D - Wow, I hope that never happens to you ND - Fears are bad conversation topics anyway, Let's move on!		D - Wow, I hope that never happens to you ND - Fears are bad conversation topics anyway, Let's move on!
Cycle 3.3	**Disclosure**	I think I'm a pretty capable being, but the other day when I was walking I fell off the table.	Open hand gestures, Head turns around the room Hand to mouth, shakes as in giggle	I think I'm a pretty capable being, but the other day I
	Prompt	What is the most embarrassing thing you did lately?	Hand gesture oriented at subject	What is the most embarrassing thing you did lately?
	SD9	D/ND		D/ND
	Closer	D - I promise to not tell anybody, ND - That's alright, I still have more to talk with you about		D - I promise to not tell anybody, ND - That's alright, I still have more to talk with you about

Cycle 3.4	**Disclosure**	I am a standard model of NAO and as you can see my head is slightly bigger than my body	Points to head with a hand	I am a standard model of Keepon and as you can see my hips are pretty big
	Prompt	Which things would you change in your appearance if you could?	Hand gesture at subject	Which things would you change in your appearance if you could?
	SD10	D/ND		D/ND
	Closer	D - I should let you know that your [BODYPART] does not make you less beautiful to me ND - You probably don't even have any		D - I should let you know that your [BODYPART] does not make you less beautiful to me ND - You probably don't even have any
Cycle 3.5	**Disclosure**	I met another NAO robot recently, got very attracted, but was too afraid to ask out on a date	Scratches head	I met another abstract robot recently, got very attracted, but was too afraid to ask out on a date
	Prompt	What are some things that you would like to do, but too scared?	Puts one arm on another, hand on mouth, as in "Hmm", inquisitively	What are some things that you would like to do, but too scared?
	SD11	D/ND		D/ND

	Closer	D - Thank you for sharing this with me! ND - No problem, you don't have to answer the questions you don't want to answer		D - Thank you for sharing this with me! ND - No problem, you don't have to answer the questions you don't want to answer
		Thank you so much for talking to me, [NAME], hope we will meet again and you tell me all about your [DREAM]	Opens arms, waves goodbye	Thank you so much for talking to me, [NAME], hope we will meet again and you tell me all about your [DREAM]

Appendix 2

Pre-experiment questionnaire

For the MS Media Technology
Leiden University
Graduation Project Experiment

1. **I am**

male □

female□

other □

2. I am _____ years old

3. Do you have prior experience of interacting with robots?

Yes □
No □

4. Describe your prior experience, if you had any?

Here are a number of personality traits that may or may not apply to you. Please write a *number* next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

Disagree strongly	**Disagree moderately**	**Disagree a little**	**Neither agree nor disagree**	**Agree a little**	**Agree moderately**	**Agree strongly**

1 2 3 4 5 6 7

I see myself as

1. _____ Extraverted, enthusiastic.
2. _____ Critical, quarrelsome
3. _____ Dependable, self-disciplined
4. _____ Anxious, easily upset
5. _____ Open to new experiences, complex
6. _____ Reserved, quiet.
7. _____ Sympathetic, warm.
8. _____ Disorganized, careless.
9. _____ Calm, emotionally stable.
10. _____ Conventional, uncreative

Appendix 3

Post-experiment questionnaire
For the MS Media Technology
Leiden University
Graduation Project Experiment

Please rate your impression of the robot on these scales

Fake	1 2 3 4 5	Natural
Machinelike	1 2 3 4 5	Humanlike
Unconscious	1 2 3 4 5	Conscious
Artificial	1 2 3 4 5	Lifelike
Dead	1 2 3 4 5	Alive
Stagnant	1 2 3 4 5	Lively
Object	1 2 3 4 5	Being

Mechanical	1	2	3	4	5	Organic
Inert	1	2	3	4	5	Interactive
Apathetic	1	2	3	4	5	Responsive
Dislike	1	2	3	4	5	Like
Unfriendly	1	2	3	4	5	Friendly
Unkind	1	2	3	4	5	Kind
Unpleasant	1	2	3	4	5	Pleasant
Awful	1	2	3	4	5	Nice

Do you feel like you would want to speak to this robot again?

Any other thoughts on the experience?
(continue on the other side of the paper)

Appendix 4

Informed Consent Form

For the MS Media Technology
Leiden University
Graduation Project Experiment

Researcher
Maral Gurbanzadeh
Advisors
Edwin van der Heide
Joost Broekens

You are being asked to take part in a research study.

Before you decide to participate in this experiment, please read the following information.

Purpose of the study

The purpose of this study is to research the effects of a robot's appearance in a conversation.

Results of this study are important for the body of research in human-robot interaction concerning privacy and protection of personal data as well as psychological effects that a robot's design may have on the behaviour of human users.

Study procedures

You are requested to take a personality test, after which you will have a conversation with a robot. Some of the questions will be personal. After the interview I will ask you to fill the questionnaire about your impressions of the meeting. Your involvement in the experiment will require approximately 20 minutes. Your responses during the interview will be recorded in audio and video format for analysis.

You may decline to answer any or all questions of the interview/questionnaires and you may also terminate your involvement at any time if you choose.

Confidentiality

Your data will be kept confidential. Your responses will be anonymized and assigned a code to be used in all research documents. Your verbal responses during the interview will be transcribed into text, anonymized and analyzed by independent judges.

Visual data with identifiable information will not be disclosed to third parties, will not be published (unless you sign a media release document), will always be stored offline, protected with a password and permanently deleted after the analysis is finished.

Please do not write any identifying information on your participation sheet.

Reward

You are invited to participate in a random prize draw among participants of this experiment to win an Amazon Echo Dot device. You will be assigned a separate number for this prize draw, the winner of which will be announced after all the participants have been recruited.

The content of your answers, decisions to not answer or to terminate your involvement in the experiment early will not affect your participation in the prize draw.

Consent

I have read and understood the provided information and have had the opportunity to ask questions. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason and without cost. I voluntarily agree to take part in this study.

Participants printed name

Participant's signature _____ Date _____

Researcher's signature _____ Date _____

If you have any questions about the study please contact Maral at 0623540073 or Edwin van der Heide at +31715277033