

Company from the Uncanny Valley: A Psychological Perspective on Social Robots, Anthropomorphism and the Introduction of Robots to Society



Janina Luise Samuel

(University of Stirling; janina.samuel@arcor.de)

ORCID: 0000-0002-3473-4761

Introduction

The introduction of social robots into society has sparked controversy, with concerns originating in ethical, legal, economical and personal grounds. Amidst a recent shift in interest in robotics onto social robots and predictions that social robots will likely become increasingly important in society (Ullrich & Diefenbach 2017, 19), the need to resolve these issues and find common ground on our understanding of these new agents in our midst only grows more pressing.

Among the main concerns in this field are the topics of how much of our social lives robots should be allowed to occupy, how and when they should be permitted to interact with vulnerable groups such as children or those in care environments, and robot humanisation, in which robots designed to appear more like people are also understood to be more like people. The former topics are debated and investigated much in relation to the possibility of introducing more social robots in elderly care, where many fear they may have negative effects on the wellbeing of the people they care for, cause isolation issues and more (Sharkey & Sharkey 2012, 18). The latter tendency has attracted much interest and concern as well, and some have made strong claims that robots should ideally never be humanised at all (Bryson 2010, 15).

This paper shall present a psychological perspective on this topic. While Psychology cannot give answers to all the wider moral and legal issues that have been raised about robot humanisation, it can at least provide an understanding as to why the tendency to humanise can seem so pervasive and present a framework in which to understand human responses to robots in general. This paper shall also attempt to apply some of this knowledge to current debates and concerns around social robots and their introduction to society and offer some ideas on how these could be approached with the psychological

perspective in mind.

I. Part One: Psychological Perspectives on Robot Humanisation

The first part of this paper will discuss how robots are perceived by humans and why they may be perceived in this way. Two key terms that will be used are anthropomorphism and humanisation. Anthropomorphism shall refer to its psychological definition, i.e. the human tendency to ascribe human-like features to nonhuman entities, while humanisation shall refer to a state in which a robot would be fully perceived as a person, not just described as having human-like traits.

1.1 Human responses to social robots

The human tendency to anthropomorphise is a widely known phenomenon and has long been a main factor in social robot design, as it would seem that robots whose design encourages anthropomorphism are accepted more easily in human-robot-interactions (HRIs) (Złotowski *et al.* 2015, 19). Because of this, many social robots are designed to have some human-like traits.

This human preference has been documented in several studies on HRI that focus on social robots' communication – a particularly important feature, considering that it is key to their social acceptance and thus ability to succeed in whatever task they were designed for. Not only do people show a general preference for spoken language over mechanical sounds (Sims *et al.* 2009, 18), but they also trust autonomous devices such as self-driving cars more when their voice has anthropomorphic features, such as a name or gender (Waytz, Heafner & Epley 2014, 19).

However, this higher acceptance for anthropomorphic robots is not universal. Instead, people show a preference for robots' design to be matched to their task. While it is important for a social robot to have human-like features, an industrial robot may be thought of in a different manner and thus does not appear to need to look human in order to be deemed acceptable for their task by a human observer (Goetz, Kiesler & Powers 2003, 16).

1.1.1 Individual differences in responses

Gender differences

There are examples of studies which have found that people of different genders respond differently to gendered robots (Crowelly *et al.* 2009, 16). The reasons behind this are not yet fully explored, but some theories suggest that there is a projection of gender expectations onto the robot when a robot is equipped with certain gender features, leading to either identification with it or a perception of it as being an opposite of oneself. This would then lead to a more positive response if the robot is perceived to be more like oneself, or

a less positive one if it is designed with features of the opposite gender (Eyssel *et al.* 2012, 16). If this were true, this would be an example of how anthropomorphism affects HRI both positively and negatively – however, the response patterns in gendered HRIs are not as consistent as this would suggest (Crowley *et al.* 2009, 16). The mechanisms behind these different responses are therefore likely more complex and require further investigation.

Cultural differences

Aside from gender differences, differences in cultural background also impact how robots are perceived. Many studies investigating this focus on differences between Western and Asian, specifically Chinese, Korean and Japanese responses. The latter are stereotypically seen as “robot enthusiast” cultures who, if this stereotype were true, might be more accepting of social robots and anthropomorphise them more. However, this idea has found little support. Studies comparing Europeans’ and Japanese’s expectations and attitudes towards robots have found that, while there are some differences, the expected “robot enthusiasm” was nowhere to be found (Haring *et al.* 2014a, 17). Instead, expectations about robots’ functionality seem to be not too dissimilar between Western and Asian cultures in most studies (Lee *et al.* 2012, 17; Haring *et al.* 2014, 17). The differences that do exist appear to relate to social convention – one culture might be more willing to allow a robot into a social scenario than another, depending on how important the scenario in question is in the participants’ different cultures.

Differences in how robots are perceived exist among European cultures as well, which again seem to originate mostly in social norms, not in wildly different expectations towards robots in general (Cortellessa *et al.* 2008, 16). A comparable pattern is found again when comparing Middle Eastern and Western participants’ perceptions (Salem, Ziadee & Sakr 2014, 18).

While people of different cultures are not as different in their general approach to robots as stereotypes may suggest, their different social conventions are still an important factor to be understood if a social robot is to be introduced successfully in different countries. This is supported by further studies: Findings on a Japanese and Egyptian test group that were each greeted by a social robot show that participants showed a much higher acceptance of the robot if it used the greetings that were customary in their cultures, and were much less comfortable around it if it used the other culture’s greeting (Trovato *et al.* 2013, 19). Similarly, another study on such cultural differences that was run with native English- and native Arab-speakers showed how the same behaviour in the same context can elicit different responses depending on the human’s cultural background: The Arab participants had a greater preference for apologising communication strategies, and thus, after both participant groups were exposed to the same communication scenarios with a social robot including an apologetic scenario, the Arab group perceived the robot to be both more polite and were more forgiving when rating its competency (Salem, Ziadee & Sakr 2014, 18).

However, truly understanding the origins and implications of such different co-

ventions proves to be complex. A 2005 study with European and Japanese participants about expectations and desires about robots found differences in general regard of what role robots should take, with Europeans taking a more utilitarian view of robots and robot interactions, while Japanese participants showed a more social perspective with more emphasis on the relationship they would want a robot to have with family and friends rather than functional features (Bartneck *et al.* 2005, 15). These differences are argued to exist due to many different factors, reaching back to historical differences in how robots were developed and adapted into society as well as differences in cultural understanding of society and what makes a person (expressed through differences such as individualistic and holistic thinking) between Western and Japanese study participants which have been explored in detail in a 2009 paper by MacDorman, Vasudevan and Ho. It considers how, in Japan, robots have had a long tradition as humanoid agents through the art of *karakuri* puppets, and how, in contrast to cultures with roots in Judeo-Christian world-views, cultures with a background in Buddhism or Confucianism view their humanity as less unique and distinct and more as part of a whole in which spirit and agency could more readily be ascribed to nonhuman entities. Both these factors are examples of possible reasons why the Japanese view of robots may be less utilitarian and more concerned with social scenarios. The paper's analysis sheds light on how complex the relationship with robots within a culture truly is and how many factors there are to be considered to understand where attitudes towards robots come from. Studies like this would seem paramount for deepening the understanding of robot perception in all cultures around the world, and it would seem advisable to carry out further analyses of robots in other cultures in this fashion (MacDorman, Vasudevan & Ho 2009, 17).

Overall, the main reasoning behind the differences in preferences and anthropomorphism between cultures appear to be different social conventions. There are also differences in design and use preferences, as well as differences in how much robots are encountered in daily life. These factors may impact overall comfort with and response to robots as well (Haring *et al.* 2014; 2014a, 17). This highlights a need for robot localisation in order to ensure that social robots are accepted in different social environments across the world.

Previous experience with robots

Different exposure to robots may lead to different approaches and perceptions of robots. This may mean that, on an international scale, how much experience a culture has with robots may influence overall national perception of them – but this exposure is also relevant on the individual level. This is exemplified well in children: early exposure to robots may contribute to lowering feelings of apprehension around robots and help build a more positive attitude. An example of this is a 2019 study in which young children's attitudes towards robots were assessed before and after interaction with a child-friendly humanoid robot. They were asked to indicate whether they thought of robots more as an agent outside their social circle, akin to an imaginary cre-

ature, or a friend or family member. After the HRI, a number of children shifted their perception of robots closer to their human social circle. This not only shows an increase in closeness perception to robots but may also indicate an increase in anthropomorphism as they shift their perception from something imaginary to something human and project an emotional bond onto the robot (Conti, Di Nuovo & Di Nuovo 2019, 16).

1.1.2 The uncanny valley

While eliciting social responses in humans is easier when the robot front of them is human-like in design, this does not mean that robots automatically become more accepted the more human they look. This may initially be the case, but human design appears to reach a point at which positive social responses turn into negative ones and robots are rejected for seeming “too human”. This phenomenon is called the uncanny valley, which describes the point at which something nonhuman has begun to look so human that the subtle differences left appear disturbing (MacDorman 2006, 17). Examples of this in HRI can be observed in interactions with androids: Their design is human-like enough so that when their expression, movement and, if present, speech do not fully match up with that of an actual human, they are reported as seeming creepy (MacDorman 2005, 17). Meanwhile, less human-like anthropomorphic robots that are clearly distinguishable from the human form may be accepted relatively easily in comparison (Greco, Anerdi & Rodriguez 2009, 17). This presents a challenge to social robot design as efforts are made to strike a balance between human-like features and dodging the uncanny valley to maximise acceptance. This is made all the more difficult by the fact that how and when the uncanny valley-effect takes hold differs between individuals, though evidence suggests that generally, the effect is triggered when the expectations of a robot struck up by its design and its actual performance do not match (Walters *et al.* 2008, 19).

A detailed study on where individual differences stem from was carried out in 2015 and identified Animal Reminder Sensitivity, Neuroticism and its Anxiety facet and Religious Fundamentalism as predictable indicators for individually different uncanny valley-perception. The study suggests that the way these factors impact the uncanny valley-effect differs: religious fundamentalism is based on societal norms (influenced in the case of this study by the belief in human distinctiveness based on a Christian worldview), while animal reminder sensitivity appears to be based more on evolved survival mechanisms. The explanation for this would be that something nonhuman may elicit fear responses to ensure survival when faced with animals in the wild, while something reminiscent of a human that does not fully fulfil human criteria may trigger disgust responses meant to protect early humans from infections or similar by for example keeping them away from corpses or other humans with disease. According to this, the uncanny valley-effect would therefore originate both in societal norms and evolved survival mechanisms. The assumption that the uncanny valley-effect has some biological basis is further supported by findings showing that the effect may not be a purely

human phenomenon, as it has been shown that even monkeys show uncanny valley-responses when faced with real and synthetic monkey faces (Steckenfinger & Ghazanfar 2009, 18). These societal and biological factors then seem to be mediated by personality-based influences, as shown by the impact of neuroticism: people who rank high on measurements of neuroticism experience fear or disgust more easily, so the uncanny valley-effect is triggered more easily for them as well (MacDorman & Entezari 2015, 17).

However, the idea of the uncanny valley effect has not been without criticism. It has for example been suggested that the uncanny valley is not unique to human representations at all, and rather elicited any time when an image activates competing representations that cannot be reconciled – this would call for further research as, if this were true, this could expand the uncanny valley-effect to nonhuman robots as well and thus cause need for it to be considered on a much broader scale in robot design than previously assumed (Ferrey, Burleigh & Fenske 2015, 16). Furthermore, it has been questioned whether the uncanny valley is a valley at all, as there is little evidence for dislike of a synthetic humanoid representation fading as it becomes “human enough” again. Instead, it has been suggested that the uncanny valley may be an uncanny cliff, which would mean that it is impossible to create a humanoid robot that is human enough to not seem creepy (Bartneck *et al.* 2007, 15). This would again have wide implications for robot design and call the usefulness of humanoid design into question entirely. Because of the lack of evidence for the originally proposed form of the uncanny valley, it has also been suggested that it is a mistake to try and categorise responses to fit the uncanny valley-theory and that it is instead necessary to completely reconsider our understanding of this phenomenon. More research to explore how the uncanny valley — if it is a valley — really presents itself and whether it really is a unique problem in anthropomorphism towards synthetic beings is therefore needed (Złotowski, Proudfoot & Bartneck 2013, 19).

1.1.3 Psychological mechanisms of anthropomorphism

The previous section has focused mostly on individual differences in robot perception and degrees of anthropomorphising. This section shall focus on exploring the underlying psychological mechanisms of and potential reasons behind the anthropomorphising mechanism.

The human predisposition to anthropomorphise has long been studied, but the absolute reasons for it as well as how and why it is triggered remain not entirely clear. An important take on the topic is that of Caporael and Heyes, whose argument entails the theory that anthropomorphism is a cognitive default that was necessary for the social evolution of the human species (Caporael & Heyes 1997, 15). By ascribing feelings and features to other beings, early humans may have been better able to find and understand allies, developing empathy and counterbalancing egoistical interests in the process. This then formed an important basis for social exchange, was thus enforced and became so integral to the human mind that it was applied to anything that one may gain some benefit from al-

lying with. This may then have been a basis to many social constructs that are still majorly influential today, such as spiritual and religious beliefs, formed out of anthropomorphism applied to the environment, founded on the idea that by performing certain rituals, one may appease the anthropomorphised surroundings to be kinder to oneself. Anthropomorphism would also have prevailed and become automatic as an explanation for all things that would otherwise have been inexplicable in the past, such as natural disasters or strange weather phenomena: It would have been easy to ascribe personality to these things to explain them, which, again, may have fed the human tendency to keep anthropomorphising.

This evolutionary perspective could also be supported by the findings of Piaget in 1929, who, in the development of his developmental stages of childhood, theorised that young children think in terms of animism, in which they believe there to be life in objects that seem to behave in ways that may look lifelike, a moving object for example. To little surprise, it has been found that children apply this animistic belief to robots as well (Beran *et al.* 2011, 15). Furthermore, outside of Piaget's theories, children as young as 9 months have been found to attribute goal-directed behaviour to observed agents even if those agents lack human features – as a robot might (Gergely & Csibra 2003, 16). A theory to explain how anthropomorphism could be present even so early in child development is presented by Spelke's core knowledge systems. These include different domains, one of which is about agents and their goal-directed behaviour. This domain enables infants from a very young age to attribute goal-directed behaviour to certain agents, and potentially anthropomorphise (Spelke & Kinzler 2007, 18). This would again show how ingrained anthropomorphism is in the human psyche, as this would assume that the drive to anthropomorphise is something humans are simply born with.

A new theory on how and when anthropomorphism is triggered is formulated in the three-factor theory (Epley, Waytz & Cacioppo 2007, 16). This theory outlines that the factors influencing whether anthropomorphism takes place are elicited agent knowledge, effectance motivation and sociality motivation. Elicited agent knowledge refers to the "accessibility and applicability of anthropocentric knowledge", i.e. whether there is knowledge to be had in oneself that can be projected, and if this knowledge is fitting to apply. Effectance motivation is defined as the "motivation to explain and understand the behaviour of other agents", i.e. there would have to be a need to anthropomorphise, such as the encountered object being novel or its actions being otherwise hard to explain. There also has to be a need to explain what is being observed at all, some motivation to reason about it in the first place. Third, sociality motivation is described as the "desire for social contact and affiliation", a need for social interaction. In the case of robots, for example, it has been shown that people who feel lonely and thus experience a greater desire for social relations anthropomorphise them more than those who do not feel lonely (Eyssel & Reich 2013, 16).

In this theory, robots might relatively easily fit the criteria for elicited agent knowledge, as a social robot would, as previously discussed, often have human-like features that could be used as a basis for making anthropocentric judgements, and, be-

ing relatively new technology, may also elicit a reasonable amount of effectance motivation, especially if the human interacting with them has little knowledge about how they work. The sociality motivation may prove to be more difficult however, as it depends on individual circumstances and appraisal. If social robot anthropomorphism is to be encouraged for acceptance reasons, then this factor would be influenced by specific robot design choices such as human-like features and gendering in order to produce a feeling of common ground that would encourage a desire to affiliate oneself with it.

It should be noted that, when the three-factor theory was tested on robots in an experiment comparing affiliation with a humanoid and a nonhumanoid robot, the humanoid robot was anthropomorphised less than the nonhumanoid one (Crowell *et al.* 2019, 16). This would seem surprising at first, considering previous findings that robots with more human-like features should elicit higher rates of anthropomorphism, and considering that the factors elicited agent knowledge and sociality motivation should seem easier to fulfil for a robot that seems more like a human. However, while the authors of this study pointed to a different explanation, scenarios like this may warrant further consideration of the uncanny valley's effects. The more humanoid a robot seems, the more easily it might elicit agent knowledge and be more appealing to be affiliated with, but at the same time, it may also not strike the necessary balance and instead fail to seem like an agent to which previously stored knowledge on humanoid agents are fully applicable and be unappealing to be affiliated with due to its disturbing appearance. Perhaps this may be worth further investigation.

1.1.4 Robot anxiety

If one discusses the human response to robots, the topic of robot anxiety should not be overlooked. While this paper has previously focused on positive or at worst uncomfortable responses to robots, this section shall examine why and when people explicitly reject robots.

Two scales have been developed to measure negative feelings towards robots: NARS, the Negative Attitude towards Robots Scale (Nomura *et al.* 2004, 18), and RAS, the Robot Anxiety Scale (Nomura *et al.* 2006a, 18). The NARS, originally a first attempt to creating the RAS, measures negative conceptions of robots through a variety of questionnaire items in the subcategories "Negative Attitude toward Situations of Interaction with Robots", "Negative Attitude toward Social Influence of Robots" and "Negative Attitude toward Emotions in Interaction with Robots". These categories are fairly descriptive and give an overview of main areas in which negative responses to robots are most likely to surface: direct interaction with them, concerns about what their impact on society may be, and a rejection of emotional bonding with robots. These areas of negative attitude are similar to the areas in which anxiety is expressed in the RAS questionnaire's subcategories, "Anxiety toward Communication Capability of Robots", "Anxiety toward Behavioural Characteristics of Robots" and "Anxiety toward Discourse with Robots". What can be seen in these categories is that it would seem a lot of negativity and fear towards robots is rooted in a concern that robots may respond to social situations and demands in un-human

ways, as illustrated well in questionnaire items such as “Robots may talk about something irrelevant during conversation”, “Conversation with robots may be inflexible” and “Robots may be unable to understand complex stories”. There is also unease due to a lack of clarity on what a robot’s capabilities and applications are (“What power robots will have”, “What robots will do”), as well as insecurity due to this lack of knowledge about how one-self would be supposed to respond if one had to interact with a robot (“How I should talk with robots”, “I may be unable to understand the contents of robots’ utterances to me”).

Many of these concerns seem to be grounded in a lack of knowledge about robots. Perhaps this is logical – robots are still new technology to many, and so to those who have little experience with them, they would seem very novel and potentially scary. Applications of this scale have shown that there are, as there was in general perception of social robots, individual differences in robot anxiety, and that the most pressing areas of concern as well as severity of concerns differ between genders and cultures (de Graaf & Allouch 2013, 16). For example, gender differences may show that women tend to be less trusting towards robots and feel more concerns about interacting with them than men (Nomura *et al.* 2006, 18), though it is not entirely clear where these differences stem from. Meanwhile, cultural differences show that not only are robots perceived differently in general, but the fears associated with them also differ between cultures. For example, Europeans may have more practical concerns when thinking about robots in daily life, such as loss of jobs. On the other hand, Asian study participants may express greater concerns in personal and social areas, which may link back to the previously discussed difference in utilitarian and more social perception (Haring *et al.* 2014, 17). Furthermore, an interesting 2009 study on implicit attitudes towards robots revealed that US participants had a stronger association between weapons and robots than Japanese participants did, which might cause greater unease towards robots as unfeeling and potentially dangerous tools (MacDorman, Vasudevan & Ho 2009, 17). This could, among other reasons, be related to different depictions of robots in Asian and European media: robots in Western media are often enemy figures bent on destruction (“Terminator”), or generally seen as action agents used for fighting (“RoboCop”, “Transformers”). Even when robots are the heroes of their story or at least sympathetic, coexistence with them is often shown as challenging or impossible and results in conflict (“Blade Runner”, “I, Robot”). Meanwhile, examples of Japanese media in which robots coexist peacefully with humans and their integration as social agents is more seamless are much easier to find (“Doraemon”, “Astro Boy”). This difference may influence the implicit understanding of what a robot is and how it interacts with people, leading to different perceptions of and fears related to robots in real life (MacDorman, Vasudevan & Ho 2009, 17).

However, while interesting to study, it is still advisable to not overemphasize these differences, as it has been shown that there are many fears cultures share in equal measure, such as fears about loss of control and human replacement (Haring *et al.* 2014, 17).

Attempts to alleviate robot anxiety have often taken the form of exposure. The-

re is evidence that even mental simulation of a robot encounter can help to decrease fears (Kuchenbrandt & Eyssel 2012, 17), however results after actual exposure are inconsistent (Nomura & Kawakami 2011, 17). Evidence overall shows that there are attitude changes, but whether these are positive or negative differs between studies and seems to depend on robot design and individual traits (de Graaf & Allouch 2013, 16).

1.1.5 Benefits of social robots

But why is it so important to work against robot anxiety? Not only are robots becoming more and more prevalent in daily life, but rejection of robots for whatever reason means that the benefits that can come of the implementation of social robots in society are rejected as well.

Prevalent among these benefits is the idea that social robots may serve well as therapy devices. Studies investigating this possibility have shown that positive effects of HRI can objectively be measured through physical calming responses (Sefidgar *et al.* 2015, 18). Robots may also serve as agents to alleviate effects of loneliness, if applied carefully, as people who feel lonely anthropomorphise more readily and may thus accept a robot as comfort, even if this should only be a temporary measure (Eyssel & Reich 2013, 16). In interactions with the elderly at care homes, social robot companions have been shown to increase wellbeing (Wada & Shibata 2007, 19) and help dementia patients by increasing their social engagement both through interactions with the robot and other patients as they engage with the robot together (Tamura *et al.* 2004, 18). Moreover, the decrease of stress levels measured in care homes after the introduction of a social robot applies not only to the elderly, but also to their carers, who, in one study, reported feeling more at ease as they trusted their charges were occupied and happily engaged with the robot (Wada *et al.* 2004, 19). Social robots may also have other therapeutic applications, including as therapy devices for children suffering from anxiety in specific situations such as visits to the dentist or mental health issues (Alemi *et al.* 2014, 15; Libin & Libin 2004, 17), therapy support for adults struggling to overcome mental health issues in daily life (Ab Aziz 2015, 15; Libin & Libin 2004, 17), or as support in health areas outside of purely psychological concerns, such as in stroke patients' recovery processes (Tapus, Țăpuș, & Matarić 2008, 19).

II. Part Two: On Current Concerns about Social Robots and Their Introduction to Society

The second part of this paper shall focus on applying the previously established psychological perspective on human perception of social robots on current debates around the introduction of social robots to society. There is an ongoing debate on how far into society social robots should be integrated, and whether anthropomorphism and potential humanising are a harmless and interesting development of responses to modern technology, or whether they are problematic and potentially dangerous. These issues shall be presented in more detail in the following and discussed in regard to

how perspectives from Psychology might be applied to ease some of these tensions.

2.1 Concerns about the introduction of social robots to society

Arguments against robot humanisation often bear some concern that doing so may devalue features that make humans unique or point out how greater attachment to robots may lead humans to inconvenience or endanger themselves for something that is not actually alive. An example for the latter would be concerns about undue attachment to robots designed for dangerous purposes, such as underground exploration or mine detection: if such a robot were anthropomorphised too much or even humanised, humans may put themselves in harm's way to protect the robot, endangering themselves and undermining the robot's purpose (Robert Jr & You 2015, 18). Others voice more complicated concerns regarding the lack of autonomy or actual mind of a robot: In companionship or therapy applications, the messages voiced by the robot are not created by an empathetic listener, but they are pre-programmed by the companies who created it, and who may have their own interests, not the human's, at heart (Giger *et al.* 2019, 16). This concern is especially important when considering robots as companions for children and other vulnerable groups. Even if the robot's creators were to create a robot that acts only ever in the human customer's benefit, some worry that an overreliance on a companion robot would lead to issues for a healthy social life, as one may feel less need to connect with real people and sacrifice time and resources to a machine that should have been invested into other humans. This may also lead to poorer decision-making, as one may begin to rely on the robot's "opinions", even though these are based on pre-programmed information and are not as fully informed and reflective as a human's own judgement would be (Bryson 2010, 15).

Many also argue that the use of social robots in care environments for the elderly is problematic: There are concerns that, once they are supplied with a social robot, the elderly may not receive the same amount of human care and social interaction anymore. They may then feel less valued and infantilised as they are handed what some may perceive as a glorified toy and told to entertain themselves with it. If a care home resident is no longer able to distinguish whether the robot is a machine or a living creature due to disease such as dementia, the argument then becomes the ethical dilemma that letting them think they are interacting with a live creature when they are not, even if this may improve wellbeing, is deceit (Sharkey & Sharkey 2012, 18). Social relationships with the elderly may change as well, as the robot recipients' children may feel more at ease leaving their parents on their own while they have the robot for company – but questions have been raised whether they should feel less difficulty in leaving them, and whether this might increase the growing problem of elderly citizens feeling isolated and lonely (Calo *et al.* 2011, 15).

And yet, despite these concerns, it seems unlikely that progress in social robot introduction or the human anthropomorphic response to them will decrease. It is undeniable that robots are becoming more and more prevalent in many parts of society, and, considering the benefits discussed earlier, they will likely be used to in-

teract with potentially vulnerable groups on a much wider scale than they are now. It is also necessary to acknowledge that anthropomorphism is a basic automatic human response, meaning that it is unlikely that robots will, at least on an individual perception level, not be humanised, and in some therapeutic applications, this may even be helpful - so the important question to be discussed is not whether social robots should be introduced in more social contexts, but how this should be done.

One important note to take away from previously cited studies is that, while anthropomorphism is an inherent mechanism often and intentionally triggered, this does not mean that people are going to blindly humanise a machine that falls short of actual humanity. While humans may form an emotional bond with their robot companions, they do prefer robots that are not meant for companionship to not look unnecessarily human-like (Goetz, Kiesler & Powers 2003, 16) and thus do seem to differentiate between robot types and how much they are interested in anthropomorphising them. In terms of the three-factor theory, this could be an example for a robot not made for social bonding falling short on factor three, sociality motivation, as people feel little need to associate with something that is not meant to be emotionally associated with. Of course, this differentiation may become muddled depending on personal factors of the human operators that may make them more susceptible to anthropomorphising the robot after all, such as experiences of loneliness (Eyssel & Reich 2013, 16). Generally, however, it does not seem to be likely that people would anthropomorphise robots not designed for social exchange so much they would put themselves in harm's way for them.

Similarly, it does not seem to be the case that people would, at least currently, be willing to hand over many of their own social responsibilities and resources to a robot. While details vary, many people do not seem to be overly enthusiastic to allow robots too closely into their social lives as of yet, and do not seem likely to place too much trust into them or allow them to replace too many of their human-to-human social interactions (Haring *et al.* 2014, 17).

However, problems that remain are concerns related to the employment of robots with children or the elderly, and how they may inadequately be used to replace human social interaction. Here, it would perhaps be advisable to adopt a policy of clear communication and education: When introducing a robot, carers and relatives should receive information about what the robot's exact capabilities are to ensure that they are aware of what they can and cannot expect the robot to be in its competency and adequacy as a social agent. They may then be better equipped to ensure that they do not overly rely on the robot as a replacement and cause harm to their relative or ward, as they have been made aware of its shortcomings and can judge more clearly how and when they can and cannot rely on it. It may also be advisable to consider legal steps to ensure that such overreliance is avoided, such as restrictions of how much time can be allocated from contact with a human carer to contact with a robotic one, or required thresholds at which robots must be programmed to encourage contact with a human agent instead of themselves. This strategy of communica-

tion could also act to ensure a smoother introduction of social robots into society overall, considering how many of the causes of negative attitudes and anxiety towards robots appear to be rooted in a lack of knowledge. Information plays an important role in managing change (Mansoor & Kamba 2010, 17), and there has been much research on how information should best be presented to be accepted as well as possible, which often highlights the effectiveness of narratives and causality – therefore, these are communication strategies that should be considered when informing about social robots (Dahlstrom 2010, 16).

There are still many issues surrounding social robots that cannot be solved immediately through this communication and regulation strategy, such as legal and ethical questions about rights of future, more highly developed social robots or liability concerns about who must answer for it if a highly autonomous robot were to cause harm or damage. And even in the present, the suggested legal limitations can likely only act as preliminary measures to keep a social balance, and more research will be needed to establish how a healthy relationship with social robots can best be achieved long-term without the need for outside intervention like legal interaction limits, which may prove to be too artificial to be integrated well and eventually be disregarded in favour of the comfort and practicality of robot use, be it wise to do so or not.

Still, for now, the establishment of clear communication between and education of all concerned parties would seem like a reasonable first step towards establishing a sensible social robot-involving future.

2.2 Summarised suggestions for future robot design and robot introduction into society

Robots should be designed in ways that match their task. Social robots benefit from anthropomorphic features and should be localised carefully in order to fit into the culture they are going to be deployed in. The introduction of robots into a new social context should always be accompanied by clear communication about their benefits and their deficits. Through this communication strategy, unnecessary harm and disappointment may be avoided and it becomes an informed choice how much one will humanise a robot or how much engagement with a robot a carer for a vulnerable person is willing to take responsibility for.

Conclusion

Social robots are on the rise and will become a more and more important part of society. There are many justified concerns around this topic, but when it comes to robot humanisation and social replacement, perhaps we need not be as worried as many fear. While people do show some automatic response in projecting anthropomorphic features onto robots, they do not mindlessly humanise and may, with the right access to information, meet their new social companions fearlessly and ascribe enough of a personality to them to enjoy their company, but also know enough about them to not rely on them when doing so would be detrimental to their, or others', social life and wellbeing. If

this balance is struck, the benefits social robots can bring to our society as a whole and their introduction on a wide scale may hopefully mark the beginning of a side-by-side, positive relationship between humans and their new robotic helpers and companions.

References

- Ab Aziz A., Ahmad F., Yusof N., Ahmad F. K., & Yusof S. A. M. 2015. "Designing a Robot-assisted Therapy for Individuals with Anxiety Traits and States." *2015 International Symposium on Agents, Multi-Agent Systems and Robotics (ISAMSR)* (pp. 98-103). IEEE.
- Alemi M., Meghdari A., Ghanbarzadeh A., Moghadam L. J., & Ghanbarzadeh A. 2014. "Effect of Utilizing a Humanoid Robot As a Therapy-Assistant in Reducing Anger, Anxiety, and Depression." *2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM)* (pp. 748-753). IEEE.
- Bartneck C., Kanda T., Ishiguro H., & Hagita N. 2007. "Is the Uncanny Valley an Uncanny Cliff?." *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 368-373). IEEE.
- Bartneck C., Nomura, T., Kanda T., Suzuki T., & Kato K. 2005. "Cultural Differences in Attitudes towards Robots." *Proc. Symposium on Robot Companions (SSAISB 2005 Convention)* (pp. 1-4).
- Beran T. N., Ramirez-Serrano A., Kuzyk R., Fior M., & Nugent S. 2011. "Understanding How Children Understand Robots: Perceived Animism in Child-Robot Interaction." *International Journal of Human-Computer Studies* 69(7-8):539-50.
- Bryson J. J. 2010. "Robots Should Be Slaves." In: Y. Wilks (Ed), *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. Amsterdam: John Benjamins Publishing Company (pp. 63-74).
- Calo C. J., Hunt-Bull N., Lewis L., & Metzler T. 2011. "Ethical Implications of Using the Paro Robot, with a Focus on Dementia Patient Care." In: *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.
- Caporael L. R. & Heyes C. M. 1997. "Why Anthropomorphize? Folk Psychology and Other Stories." In: R. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), *Anthropomorphism, Anecdotes, and Animals*. New York: Suny Press (pp. 59-73).
- Conti D., Di Nuovo S., & Di Nuovo A. 2019. "Kindergarten Children Attitude Towards Humanoid Robots: What Is the Effect of the First Experience?." *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 630-631). IEEE.
- Cortellessa G., Scopelliti M., Tiberio L., Svedberg G. K., Loutfi A., & Pecora F. 2008. "A Cross-Cultural Evaluation of Domestic Assistive Robots." *AAAI Fall Symposium: AI*

- in Eldercare: New Solutions to Old Problems* (pp. 24-31).
- Crowell C. R., Deska J. C., Villano M., Zenk J., & Roddy Jr J. T. 2019. "Anthropomorphism of Robots: Study of Appearance and Agency." *JMIR Human Factors* 6(2):e12629.
- Crowley C. R., Villanoy M., Scheutzz M., & Schermerhornz P. 2009. "Gendered Voice and Robot Entities: Perceptions and Reactions of Male and Female Subjects." *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3735-3741). IEEE.
- Dahlstrom M. F. 2010. "The Role of Causality in Information Acceptance in Narratives: An Example from Science Communication." *Communication Research* 37(6):857-75.
- de Graaf M. M. & Allouch S. B. 2013. "The Relation between People's Attitude and Anxiety towards Robots in Human-Robot Interaction." *2013 IEEE RO-MAN* (pp. 632-637). IEEE.
- Epley N., Waytz A., & Cacioppo J. T. 2007. "On Seeing Human: A Three-Factor Theory of Anthropomorphism." *Psychological Review* 114(4):864.
- Eyssel F. & Reich N. 2013. "Loneliness Makes the Heart Grow Fonder (of Robots)—On the Effects of Loneliness on Psychological Anthropomorphism." *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 121-122). IEEE.
- Eyssel F., De Ruitter L., Kuchenbrandt D., Bobinger S., & Hegel F. 2012. "»If You Sound Like Me, You Must Be More Human«: On the Interplay of Robot and User Features on Human-Robot Acceptance and Anthropomorphism." *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 125-126). IEEE.
- Ferrey A. E., Burleigh T. J., & Fenske M. J. 2015. "Stimulus-category Competition, Inhibition, and Affective Devaluation: A Novel Account of the Uncanny Valley." *Frontiers in Psychology* 6:249.
- Gergely G. & Csibra G. 2003. "Teleological Reasoning in Infancy: The Naive Theory of Rational Action." *Trends in Cognitive Sciences* 7(7):287-92.
- Giger J. C., Piçarra N., Alves-Oliveira P., Oliveira R., & Arriaga P. 2019. "Humanization of Robots: Is It Really Such a Good Idea?." *Human Behavior and Emerging Technologies* 1(2):111-23.
- Goetz J., Kiesler S., & Powers A. 2003. "Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation." *The 12th IEEE International Workshop on Robot and Human Interactive Communication Proceedings. RO-MAN 2003* (pp. 55-60). IEEE.
- Greco A., Anerdi G. & Rodriguez G. 2009. "Acceptance of an Animaloid Robot As a Starting Point for Cognitive Stimulators Supporting Elders with Cognitive Impairments." *Revue d'Intelligence Artificielle* 23(4):523-37.
- Haring K. S., Mougnot C., Ono F., & Watanabe K. 2014. "Cultural Differences in Perception and Attitude towards Robots." *International Journal of Affective Engineering*

- 13(3):149-57.
- Haring K. S., Silvera-Tawil D., Matsumoto Y., Velonaki M., & Watanabe K. 2014a. "Perception of an Android Robot in Japan and Australia: A Cross-cultural Comparison." *International Conference on Social Robotics* (pp. 166-175). Cham: Springer.
- Kuchenbrandt D. & Eyssel F. 2012. "The Mental Simulation of a Human-Robot Interaction: Positive Effects on Attitudes and Anxiety Toward Robots." *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication* (pp. 463-468). IEEE.
- Lee H. R., Sung J., Šabanović S., & Han J. 2012. "Cultural Design of Domestic Robots: A Study of User Expectations in Korea and the United States." *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication* (pp. 803-808). IEEE.
- Libin A. V. & Libin E. V. 2004. "Person-Robot Interactions from the Robopsychologists' Point of View: The Robotic Psychology and Robotherapy Approach." *Proceedings of the IEEE* 92(11):1789-1803.
- MacDorman K. F. 2005. "Androids As an Experimental Apparatus: Why Is There an Uncanny Valley and Can We Exploit It." *CogSci-2005 Workshop: Toward Social Mechanisms of Android Science* Vol. 106118.
- MacDorman K. F. 2006. "Subjective Ratings of Robot Video Clips for Human Likeness, Familiarity, and Eeriness: An Exploration of the Uncanny Valley." *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science* (pp. 26-29).
- MacDorman K. F. & Entezari S. O. 2015. "Individual Differences Predict Sensitivity to the Uncanny Valley." *Interaction Studies* 16(2):141-72.
- MacDorman K. F., Vasudevan S. K., & Ho C. C. 2009. "Does Japan Really Have Robot Mania? Comparing Attitudes by Implicit and Explicit Measures." *AI & Society* 23(4):485-510.
- Mansoor Y. & Kamba M. A. 2010. "Information Acceptance and ICT Resistance: Promoting the Role of Information in Rural Community Development." *Library Philosophy and Practice*, 1.
- Nomura T. & Kawakami K. 2011. "Relationships between Robot's Self-Disclosures and Human's Anxiety toward Robots." *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology* 03:66-69. IEEE Computer Society.
- Nomura T., Kanda T., Suzuki T., & Kato K. 2004. "Psychology in Human-Robot Communication: An Attempt through Investigation of Negative Attitudes and Anxiety toward Robots." In: *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication* (IEEE Catalogue No. 04TH8759) (pp. 35-40). IEEE.
- Nomura T., Suzuki T., Kanda T., & Kato K. 2006. "Altered Attitudes of People toward Robots: Investigation through the Negative Attitudes toward Robots Scale."

- Proc. AAAI-06 Workshop on Human Implications of Human-Robot Interaction* Vol. 2006:29-35.
- Nomura T., Suzuki T., Kanda T., & Kato K. 2006a. "Measurement of Anxiety toward Robots." *RO-MAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 372-377). IEEE.
- Robert Jr L. P. & You S. 2015. "Subgroup Formation in Teams Working with Robots." *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2097-2102). ACM.
- Salem M., Ziadee M., & Sakr M. 2014. "Marhaba, How May I Help You? Effects of Politeness and Culture on Robot Acceptance and Anthropomorphization." In: *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 74-81). IEEE.
- Sefidgar Y. S., MacLean K. E., Yohanan S., Van der Loos H. M., Croft E. A., & Garland E. J. 2015. "Design and Evaluation of a Touch-centered Calming Interaction with a Social Robot." *IEEE Transactions on Affective Computing* 7(2):108-21.
- Sharkey A. & Sharkey N. 2012. "Granny and the Robots: Ethical Issues in Robot Care for the Elderly." *Ethics and Information Technology* 14(1):27-40.
- Sims V. K., Chin M. G., Lum H. C., Upham-Ellis L., Ballion T., & Lagattuta N. C. 2009. "Robots' Auditory Cues Are Subject to Anthropomorphism." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 53(18):1418-21. Sage CA: Los Angeles, CA: SAGE Publications.
- Spelke E. S. & Kinzler K. D. 2007. "Core Knowledge." *Developmental Science* 10(1):89-96.
- Steckenfinger S. A. & Ghazanfar A. A. 2009. "Monkey Visual Behavior Falls into the Uncanny Valley." *Proceedings of the National Academy of Sciences* 106(43):18362-6.
- Tamura T., Yonemitsu S., Itoh A., Oikawa D., Kawakami A., Higashi Y., ... & Nakajima K. 2004. "Is an Entertainment Robot Useful in the Care of Elderly People with Severe Dementia?." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 59(1):M83-M85.
- Tapus A., Țăpuș C., & Matarić M. J. 2008. "User—Robot Personality Matching and Assistive Robot Behavior Adaptation for Post-stroke Rehabilitation Therapy." *Intelligent Service Robotics* 1(2):169.
- Trovato G., Zecca M., Sessa S., Jamone L., Ham J., Hashimoto K., & Takanishi A. 2013. "Cross-cultural Study on Human-Robot Greeting Interaction: Acceptance and Discomfort by Egyptians and Japanese." *Paladyn, Journal of Behavioral Robotics* 4(2):83-93.
- Ullrich D. & Diefenbach S. 2017. "Truly Social Robots-Understanding Human-Robot Interaction from the Perspective of Social Psychology." In: *VISIGRAPP (2: HUCAPP)* (pp. 39-45).
- Wada K. & Shibata T. 2007. "Living with Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House." *IEEE Transactions on Robotics*

23(5):972-980.

- Wada K., Shibata T., Saito T., & Tanie K. 2004. "Effects of Robot-assisted Activity for Elderly People and Nurses at a Day Service Center." *Proceedings of the IEEE* 92(11):1780-8.
- Walters M. L., Syrdal D. S., Dautenhahn K., Te Boekhorst R., & Koay K. L. 2008. "Avoiding the Uncanny Valley: Robot Appearance, Personality and Consistency of Behavior in an Attention-seeking Home Scenario for a Robot Companion." *Autonomous Robots* 24(2):159-78.
- Waytz A., Heafner J. & Epley N. 2014. "The Mind in the Machine: Anthropomorphism Increases Trust in an Autonomous Vehicle." *Journal of Experimental Social Psychology* 52:113-7.
- Złotowski J., Proudfoot D., & Bartneck C. 2013. "More Human Than Human: Does the Uncanny Curve Really Matter?." *Proceedings of the HRI 2013 Workshop on Design and Humanlikeness in HRI from Uncanny Valley to Minimal Design Tokyo* (pp. 7-13).
- Złotowski J., Proudfoot D., Yogeewaran K., & Bartneck C. 2015. "Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction." *International Journal of Social Robotics* 7(3):347-60.

Janina Luise Samuel (Stirling)

**Company from the Uncanny Valley: A Psychological Perspective on Social Robots,
Anthropomorphism and the Introduction of Robots to Society**

Abstract: There are many issues surrounding the introduction of social robots into society, including concerns about how they may be used to replace true social interaction in personal life, dehumanise formerly social occupations such as elderly care, and be perceived as more human than they actually are. This paper shall present a psychological perspective on the human reception of social robots and apply the gathered information to address these concerns.

Keywords: social robots; anthropomorphism; robot psychology; uncanny valley.

Ethics in Progress (ISSN 2084-9257). Vol. 10 (2019). No. 2, Art. #2, pp. 8-26.

Creative Commons BY-SA 4.0

DOI:10.14746/eip.2019.2.2