CrossMark

Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction

**Jakub Złotowski · Diane Proudfoot ·
Kumar Yogeeswaran · Christoph Bartneck**

Abstract Anthropomorphism is a phenomenon that
describes the human tendency to see human-like shapes in the
environment. It has considerable consequences for people's
choices and beliefs. With the increased presence of robots,
it is important to investigate the optimal design for this tech-
nology. In this paper we discuss the potential benefits and
challenges of building anthropomorphic robots, from both a
philosophical perspective and from the viewpoint of empir-
ical research in the fields of human–robot interaction and
social psychology. We believe that this broad investigation
of anthropomorphism will not only help us to understand the
phenomenon better, but can also indicate solutions for facil-
itating the integration of human-like machines in the real
world.

J. Złotowski (✉) · C. Bartneck
HIT Lab NZ, University of Canterbury, Christchurch,
New Zealand
e-mail: jakub.zlotowski@pg.canterbury.ac.nz

C. Bartneck
e-mail: christoph.bartneck@canterbury.ac.nz

D. Proudfoot
Department of Philosophy, University of Canterbury,
Christchurch, New Zealand
e-mail: diane.proudfoot@canterbury.ac.nz

K. Yogeeswaran
Department of Psychology, University of Canterbury,
Christchurch, New Zealand
e-mail: kumar.yogeeswaran@canterbury.ac.nz

1 Introduction

Anthropomorphism has a major impact on human behav-
iour, choices, or laws. Based on evidence for the presence
of mind, captive chimpanzees were granted limited human
rights in Spain [2]. Moreover, people often refer to our
planet as 'Mother Earth' and anthropomorphism is often
used in discussions regarding environmental issues [151].
People regularly make anthropomorphic attributions when
describing their surrounding environment including animals
[44], moving geometrical figures [82] or weather patterns
[77]. Building on this common tendency, anthropomorphic
form has been used to design technology [52] and sell
products [1,4].

Anthropomorphic design is an especially important topic
for design of robots due to their higher anthropomorphizabil-
ity [91]. Increasing number of industrial and service robots
yields a question of designing this technology in order to
increase its efficiency and effectiveness. One of the main
themes in the field of human–robot interaction (HRI) tries
to address that issue. However, taking a broader perspective
that involves related fields could foster the discussion. In this
paper we present viewpoints of empirical work from HRI
and social psychology, and a philosophical discourse on the
issue of designing anthropomorphic technology.

In the next section of this paper we present perspectives
from different fields on the process of anthropomorphization
and how the general public's view differs from the scientific
knowledge. In Sect. 3 we include a broad literature review of
research on anthropomorphism in the field of HRI. In Sect. 4
we discuss why creating human-like robots can be beneficial
and what opportunities it creates for HRI. Section 5 is dedi-
cated to a discussion on potential problems that human-like
technology might elicit. In Sect. 6 we propose solutions to
some of these problems that could be applied in HRI.

2 Why Do We Anthropomorphize?

From a psychological perspective, the central questions about anthropomorphism are: what explains the origin and persistence of anthropomorphism? Psychologists and anthropologists have explained the origin of anthropomorphism as an adaptive trait, for example with respect to theistic religions. They speculate that early hominids who interpreted ambiguous shapes as faces or bodies improved their genetic fitness, by making alliances with neighbouring tribes or by avoiding threatening neighbours and predatory animals [10,22–24,27,75]. But what explains the persistence of anthropomorphizing? Here theorists hypothesize that there are neural correlates of anthropomorphizing [123,138], specific anthropomorphizing mechanisms (e.g. the hypothesized hypersensitive agency detection device or HADD [11,12]), and diverse psychological traits that generate anthropomorphizing behaviour [28]. They also suggest (again in the case of theistic religions) that confirmation bias [10] or difficulties in challenging anthropomorphic interpretations of the environment may underlie the persistence of anthropomorphizing [115].

Epley et al. [53] proposed a theory that determined three psychological factors as affecting when people anthropomorphize non-human agents:

1. Elicited agent knowledge—due to people's much richer knowledge regarding humans compared with non-human agents, people are more likely to use anthropomorphic explanations of non-human agents' actions until they create an adequate mental model of non-humans.
2. Effectance motivation—when people are motivated to explain or understand an agent's behaviour the tendency to anthropomorphize increases.
3. Sociality motivation—people who lack social connection with other humans often compensate for this by treating non-human agents as if they were human-like [54].

This theory has been also successfully applied in the context of HRI [56].

Although people make anthropomorphic attributions to various types of non-human agents, not all agents are anthropmorphized in the same way. Anthropomorphization of animals is distinct from the tendency to anthropomorphize artifacts, such as cars or computers [40]. There are gender differences in this tendency when the target is an animal, with females more likely to make anthropomorphic attributions than males. However, when anthropomorphizing machines, males and females are equally likely to exhibit this tendency.

2.1 A Philosophical Perspective on Anthropomorphism

From a philosophical perspective, the central questions about anthropomorphism are: can we make a principled distinction between justified and unjustified anthropomorphism, and if so how? Is anthropomorphizing (philosophically) justified? If anthropomorphism is a 'natural' behaviour, this question may seem odd. Moreover, some researchers in AI have argued that there is not anything to the notion of mind but an evolved human tendency to anthropomorphize; an entity's having a 'soul' is nothing over and above the tendency of observers to see the entity in this way. On this view, humans are 'natural-born dualists' [26, p. xiii]. However, the intuition that anthropomorphizing can be illicit is strong—for example, my automobile is not a person, even if I attribute human and personal characteristics (and have affective responses) to it. Researchers in AI claim that their machines have cognitive and affective characteristics, either explicitly or implicitly, and this is in particular true in the case of anthropomorphic robots (and anthropomorphic software agents). These assertions require philosophical analysis and evaluation, along with empirical investigation. Determining under what conditions the anthropomorphizing of machines is justified and under what conditions unjustified is central to the question whether 'expressive' or 'emotional' robots actually have emotions (see e.g. [3,6]). It is also key to the growing debates within AI about the ethical use of artificial systems.

In this regard, a combination of philosophical analysis and experimental work is required, but to our knowledge has not been carried out. In the large body of experimental work on human reactions to anthropomorphic robots, responses on standard questionnaires are commonly taken to demonstrate that subjects identify a robot's displays or movements as (for example) expressions of the fundamental human emotions—happiness, sadness, disgust, and so on (see e.g. [34]). The robot is said to smile or frown. However, taking these responses (in forced choices) at face value ignores the possibility that they are elliptical for the subjects' actual views. To use an analogy, it is common when discussing fictions to omit the logical prefixes such as 'I imagine that ...' or 'Make-believedly ...'—for example, we say 'Sherlock Holmes lives at 221B Baker Street' when we mean 'In the fiction Sherlock Holmes lives at 221B Baker Street'. Something similar may be occurring in discussions of anthropomorphic robots; saying that the robot has a 'happy' expression might be shorthand for the claim (for example) that *if* the robot were a human, it *would have* a happy expression. A fine-grained 'philosophical' experiment might allow us to find out if this is the case. Experimental philosophy has gained ground in some areas of traditional a priori argument such as ethics; it might be used in AI to enable more accurate analysis of human reactions to anthropomorphic robots.

2.2 Science-fiction as a Proxy to General Public's Perception

It can be argued that almost all prior knowledge of participants about robots in HRI studies stem from the media. An extensive discussion on how robots are being portrayed in the media is available [14]. Here therefore only in short: there are two main story types that run through the media about robots. One is that robots want to be like humans (e.g. Mr. Data) and that once a superior level of intelligence and power is achieved will want to kill or enslave humanity (e.g. Lore). These rather negative views on the future of human–robot relationships are based on the media industry's need to produce engaging stories. Fear is the single most used method to engage the audience. A future world in which humans and robots live happily side by side is rare. The TV show Futurama and the movie Robot and Frank comes to mind as the glowing exceptions. The stories presented in the media that focus on robots can be categorized along the questions whether the the body and/or the mind of the robot is similar to humans. If we take Mr. Data again as an example, he does look very much like a human, but his mind functions differently. From this the writers can form engaging themes, such as Data's quest to understand humor and emotions. And we are surprised when Data emerges from the bottom of a lake without any specific gear. His highly human-like form makes us believe that he might also need oxygen, which he does not. In summary, the media has used robots extensively and most of the knowledge and expectations that the people on the street have are based on these media and not on the scientific literature.

3 Anthropomorphism in HRI

Reeves and Nass [116], in their classical work on the 'Computers are Social Actors' paradigm, showed that people engage in social interaction with various types of media. Therefore, designers of interactive technologies could improve this interaction, building on the chronic tendency of people to anthropomorphize their environment. Due to their higher anthropomorphizability and physical autonomy in a natural human environment, robots are especially well suited to benefit from anthropomorphism [91]. Furthermore, physical presence in the real world (rather than being merely virtual) is an important factor that can also increase the anthropomorphic quality of robots [92]. Mere presence of a robot was found to lead to the social facilitation effect [120]. Moreover, when playing a game against a robotic opponent, people may utilize similar strategies as when they play against a human [137]. They also hold robots more accountable for their actions than other non-human objects [87]. This tendency cannot be observed when an opponent is a disembodied computer. On the other hand, Levin et al. [95] suggests that people initially equate robots and disembodied computers in terms of intentionality. However, when they focus on the intentional behaviour of a robot, this tendency can be overridden.

3.1 Factors Affecting Anthropomorphism

It is important to remember that anthropomorphism is affected not only by physical appearance. Hegel et al. [81] created a typology of signals and cues that robots emit during interaction and which can affect their perceived human-likeness. Choi and Kim [41] proposed that anthropomorphism of robots involves: appearance, human–robot interaction, and the accordance of the two former measurements. The distinction between the anthropomorphic form in appearance and in behaviour can also be found in the model presented by von Zitzewitz et al. [155].

External appearance can influence the perception of an object [129]. According to Fong et al. [65], we can classify robots based on their appearance into four categories: anthropomorphic, zoomorphic, caricatured, and functional. In the field of robotics there is an increased tendency to build robots that resemble humans in their appearance. In recent years we can observe an increased number of robots that are built with legs rather than wheels [39]. Some researchers suggest that, in order to create robots with an anthropomorphic appearance that are capable of engaging in interaction with humans in a way analogous to human-human interaction, it is necessary to build robots with features that enable them to perceive the world similarly to humans, i.e. using two cameras (in place of eyes) and two microphones (ears) [133]. Di Salvo et al. [49] state that it is the presence of certain features and the dimensions of the head that have a major impact on the perception of a humanoid's head as human-like (Fig. 1). Anthropomorphic form in appearance has even been attributed to flying robots [43].

However, research into anthropomorphism in the field of HRI has not been limited to the anthropomorphic form of a robot's appearance. HRI factors were found to be even more important than embodiment in the perceived humanness of robots [90]. Kahn et al. [86] presented six benchmarks in HRI that constitute essential features affecting robots' perceived human-likeness: autonomy, imitation, intrinsic moral value, moral accountability, privacy, and reciprocity. Previous studies proposed other factors, such as verbal [132] and non-verbal [100,122] communication, the perceived 'emotions' of the robot [59], the intelligence of the machine [15] or its predictability [60]. Moreover, robots that exhibit typically human behaviours, such as cheating, are also perceived as more human-like [131]. There is a philosophical question, whether such behaviour really makes robots more human-like or if instead it is necessary for them to 'truly' feel
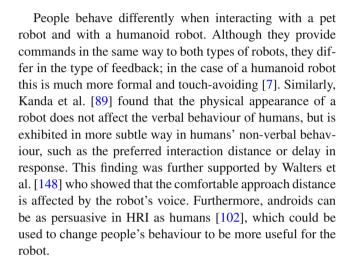
Fig. 1 Robots with different anthropomorphic features in appearance. From the *left:* Telenoid, Robovie R2, Geminoid HI2, Papero, NAO

emotions and have intentions. However, Turkle [144] points out that the behaviour of robots is more important than their inner states for them to be treated as companions.

Furthermore, anthropomorphism is the result not only of a robot's actions, but also of an observer's characteristics, such as motivation [53], social background [55], gender [61], and age [88]. Moreover, the social relationship between a robot and a human can affect the degree to which a robot is attributed humanness. People apply social categorizations to robots and those machines that are perceived as ingroup members are also anthropomorphized more strongly than outgroup robots [57,93]. Therefore, it should not be surprising that a robot that has the physical appearance of a member of another race is treated as a member of an outgroup and perceived as less human-like by people with racial prejudices [58]. There is also empirical evidence that mere HRI can lead to increased anthropomorphization of a robot.

3.2 Consequences of Anthropomorphizing Robots

Despite multiple ways in which we can make robots more human-like, anthropomorphism should not be a goal on its own. People differ in their preferences regarding the appearance of a robot. These differences can have cultural [55] or individual (personality) [134] origins. Goetz et al. [70] emphasized that, rather than aiming to create the most human-like robots, embodiment should be designed in a way that matches the robots' tasks. Anthropomorphism has multiple desirable and undesirable consequences for HRI. A robot's embodiment affects the perception of their intelligence and intentionality on neuropsychological and behavioural levels [80]. More visual attention is attracted by anthropomorphic or zoomorphic than inanimate robots [9]. Furthermore, similar perceptual processes are involved when observing the movement of a humanoid and of a human [104]. Based on the physical appearance of a robot, people attribute different personality traits to it [149]. Moreover, people use cues, such as a robot's origin or the language that it speaks, in order to create a mental model of the robot's mind [94].

People behave differently when interacting with a pet robot and with a humanoid robot. Although they provide commands in the same way to both types of robots, they differ in the type of feedback; in the case of a humanoid robot this is much more formal and touch-avoiding [7]. Similarly, Kanda et al. [89] found that the physical appearance of a robot does not affect the verbal behaviour of humans, but is exhibited in more subtle way in humans' non-verbal behaviour, such as the preferred interaction distance or delay in response. This finding was further supported by Walters et al. [148] who showed that the comfortable approach distance is affected by the robot's voice. Furthermore, androids can be as persuasive in HRI as humans [102], which could be used to change people's behaviour to be more useful for the robot.

4 Benefits and Opportunities of Anthropomorphic Robots

From the literature review presented in the previous section, it becomes clear that there are multiple ways in which robots can be designed in order to create an impression of human-likeness. This creates an opportunity to positively impact HRI by building on the phenomenon of anthropomorphism. DiSalvo and Gemperle [48] suggested that the four main reasons for designing objects with anthropomorphic shapes are: keeping things the same (objects which historically had anthropomorphic forms maintain this appearance as a convention), explaining the unknown (anthropomorphic shapes can help to explain products with new functionality), reflecting product attributes (using anthropomorphic shapes to emphasize a product's attributes) and projecting human values (influencing the experience of a product via the socio-cultural context of the user).

4.1 Facilitation of HRI

The practical advantage of building anthropomorphic robots is that it facilitates human-machine interaction (see e.g.

[62,63,69,147]). It also creates familiarity with a robotic system [41] and builds on established human skills, developed in social human-human interactions [129]. A human-like machine enables an untrained human user to understand and predict the machine's behaviour—animatronic toys and entertainment robots are an obvious example but anthropomorphizing is valuable too in the case of industrial robots (see e.g. Rod Brooks's Baxter).[1] Believability is particularly important in socially assistive robotics (see [136]). In addition, where a machine requires individualized training, a human-like appearance encourages human observers to interact with the machine and so produces more training opportunities than might otherwise be available [32].

Considering that social robots may be used in public spaces in the future, there is a need for ensuring that people will treat them properly, i.e. not destroy them in an act of vandalism. We already know that people are less reluctant to punish robots than human beings [16], although another study did not show any difference in punishment between a dog-like robot AIBO and a human partner [118]. Furthermore, lighter, but nevertheless still negative, forms of abuse and impoliteness towards a robot can occur when a robot is placed in a social environment [117]. Therefore, it is necessary to counteract these negative behaviours towards a robot. Anthropomorphism could be used in order to increase people's willingness to care about the well-being of robots. Robots that are human-like in both appearance and behaviour are treated less harshly than machine-like robots [17,20]. This could be related to higher empathy expressed towards anthropomorphic robots, as their appearance and behaviour can facilitate the process of relating to them [119]. A robot that expresses 'emotions' could also be treated as more human-like [59], which could change people's behaviour.

Depending on a robot's task, different levels of anthropomorphism might be required. A robot's embodiment affects the perception of the robot as an interaction partner [64]. The physical appearance of the robot is often used to judge its knowledge [109]. Therefore, by manipulating the robot's appearance it is possible to elicit different levels of information from people, i.e. less if conversation efficiency is desired or more when the robot should receive a robust amount of detailed feedback. Furthermore, people comply more with a robot whose degree of anthropomorphism matches the level of a task's seriousness [70]. In the context of educational robots that are used to teach human pupils, a robot that can employ social supportive behaviour while teaching can lead to superior performance by students [121]. Moreover, in some cul-

tures anthropomorphic robots are preferred over mechanistic robots [13].

4.2 Anthropomorphism as a Psychological Test-bed

From a psychological perspective, human-like robots present a way to test theories of psychological and social development. It may be possible to investigate hypotheses about the acquisition (or deficit) of cognition and affect, in particular the development of theory of mind (TOM) abilities ([29,71,125,126]), by modeling the relevant behaviours on robots (e.g. [127]). Doing so would enable psychological theories to be tested in controlled, standardized conditions, without (it is assumed) ethical problems regarding consent and treatment of infant human subjects. Here practical and theoretical research goals are linked: devices such as robot physiotherapists must be able to identify their human clients' interests and feelings and to respond appropriately—and so research on the acquisition of TOM abilities is essential to building effective service robots.

4.3 Philosophical Origins of Human-like Robots

From a philosophical perspective, two striking ideas appear in the AI literature on anthropomorphic robots. *First*, the notion of building a socially intelligent robot (see e.g. [29,33]). This replaces AI's grand aim of building a human-level intelligent machine (or Artificial General Intelligence (AGI)) with the language and intellectual abilities of a typical human adult—a project that, despite some extravagant claims in the 1980s, has not succeeded. Instead the (still-grand) goal is to construct a machine that can interact with human beings or other machines, responding to normal social cues. A notable part of this is the aim to build a machine with the cognitive and affective capacities of a typical human infant (see e.g. [128,130]). For several researchers in social and developmental robotics, this involves building anthropomorphic machines [85]. *Second*, the notion that there is nothing more to the development of intentionality than anthropomorphism. We unwittingly anthropomorphize human infants; a carer interprets a baby's merely reflexive behaviour as (say) social smiling, and by smiling in return encourages the development of social intelligence in the baby. Some robotics researchers suggest that, if human observers interact with machines in ways analogous to this carer-infant exchange, the result will be intelligent machines (see e.g. [29,36]). Such interaction will be easiest, it is implied, if it involves anthropomorphic robots. The combination of these two ideas gives AI a new take on the project of building a thinking machine.

Many concepts at the centre of this work in current AI were present at the birth of the field, in Turing's writings on machine intelligence. Turing [139,141] theorized that the cortex of the human infant is a learning machine, to be orga-

---

[1] See http://www.rethinkrobotics.com/resources/videos/.

nized by a suitable process of education (to become a universal machine) and that simulating this process is the route to a thinking machine. He described the child machine, a machine that is to learn as human infants learn (see [113]). Turing also emphasized the importance of embodiment, particularly human-like embodiment [139,141]—he did, though, warn against the (hypothesized) uncanny valley [99] (see section below), saying that too human-like machines would have 'something like the unpleasant quality of artificial flowers' [143, p. 486]. For Turing, the project of building a child machine has both psychological and philosophical benefits. Concerning the former, attempting to construct a thinking machine will help us, he said, to find out how human beings think [140, p. 486]. Concerning the latter, for Turing the concept of the child machine is inextricably connected to the idea of a genuine thinking thing: the machine that learns to generalize from past education can properly be said to have 'initiative' and to make 'choices' and 'decisions', and so can be regarded as intelligent rather than a mere automaton [141, p. 429], [142, p. 393].

5 Disadvantages of Anthropomorphism in HRI

Despite numerous advantages that anthropomorphism brings to HRI, there are also some drawbacks related to human-like design and task performance. Anthropomorphism is not a solution, but a mean of facilitating an interaction [52]. When a robot's human-likeness has a negative effect on interaction, it should be avoided. For example, during medical check-ups conducted by a robot, patients felt less embarrassed with a machine-like robot than a more humanoid robot [20]. Furthermore, the physical presence of a robot results in a decreased willingness of people to disclose an undesirable behaviour compared to a projected robot [92]. These findings suggest that a machine-like form could be beneficial, as patients might provide additional information—which otherwise they might have tried to hide, if they thought it embarrassing—that could help toward a correct diagnosis. Moreover, providing an anthropomorphic form to a robot might not be sufficient to facilitate people's interaction with it. People engage more in HRI when it is goal-oriented rather than in a pure social interaction [8].

Furthermore, a robot's anthropomorphism leads to different expectations regarding its capabilities and behaviour compared to machine-like robots. People expect that human-like robots will follow human social norms [135]. Therefore, a robot that does not have the required capabilities to do so can decrease the satisfaction of their human partners in HRI. Although in the short term this can be counter-balanced by the higher reward-value due to the robot's anthropomorphic appearance [135]. In the context of search and rescue, people felt calmer when a robot had a non-anthropomorphic

appearance; considering that such an interaction context is highly stressful for humans, apparently a robot's machine-like-aspects are more desirable [25]. A similar preference was shown in the case of robots that are designed to interact in crowded urban environments [67]. People indicate that in the first place a robot should be functional and able to complete its tasks correctly and only in the second place does its enjoyable behaviour matter [96]. In addition, a robot's movement does not need to be natural because in some contexts people may prefer caricatured and exaggerated behaviour [150]. There are also legal questions regarding anthropomorphic technology that must be addressed. Android science has to resolve the issue of the moral status of androids—or unexpected ramifications might hamper the field in the future [37].

5.1 The Risks of Anthropomorphism in AI

The disadvantages of building anthropomorphic robots also include the following, in ascending order of seriousness for AI. First, it has been claimed, the phenomenon of anthropomorphic robots (at least as portrayed in fiction) encourages the general public to think that AI has advanced further than it has in reality—and to misidentify AI as concerned only with human-like systems. Jordan Pollack remarked, for example, 'We cannot seem to convince the public that humanoids and Terminators are just Hollywood special effects, as science-fictional as the little green men from Mars!' [108, p. 50]. People imagine that service robots of the future will resemble robots from literature and movies [101].

The second problem arises specifically for those researchers in social and developmental robotics whose aim is to build anthropomorphic robots with the cognitive or affective capacities of the human infant. Several theorists claim that focusing on human-level and human-like AI is a hindrance to progress in AI: research should focus on the 'generic' concept of intelligence, or on mindless intelligence [108], rather than on the parochial goal of human intelligence (see e.g. [66,79]). To quote Pollack again, AI behaves 'as if human intelligence is next to godliness' [108, p. 51]. In addition, critics say, the difficult goal of human-like AI sets an unrealistic standard for researchers (e.g. [42]). If sound, such objections would apply to the project of building a child machine.

The third difficulty arises for any attempt to build a socially intelligent robot. This is the forensic problem of anthropomorphism – the problem of how we are reliably able to detect intelligence in machines, given that the tendency to anthropomorphize leads us to find intelligence almost everywhere [110,112]. Researchers in AI have long anthropomorphized their machines and anthropomorphic robots can prompt fantasy and make-believe in observers and researchers alike.

Such anthropomorphizing is not 'innocent': instead it introduces a bias into judgements of intelligence in machines and so renders these judgements suspect.[2] Even at the beginning of the field, in 1948, Turing said that playing chess against a 'paper' machine (i.e. a simulation of machine behaviour by a human being using paper and pencil) 'gives a definite feeling that one is pitting one's wits against something alive' [141, p. 412]. His descriptions of his own machines were sometimes extravagantly anthropomorphic—he said, for example, that his child machine could not be sent to school 'without the other children making excessive fun of it' [139, pp. 460-1]—but they were also plainly tongue-in-cheek. He made it clear, when talking of 'emotional' communication between human and child-machine[3] (the machine was to be organised by means of 'pain' and 'pleasure' inputs) that this did 'not presuppose any feelings on the part of the machine' [139], [141, p. 461]. In Turing's vocabulary, 'pain' is just the term for a signal that cancels an instruction in the machine's table.

Anthropomorphizing leaves AI with no trustworthy way of testing for intelligence in artificial systems. At best, the anthropomorphizing of machines obscures both AI's actual achievements and how far it has to go in order to produce genuinely intelligent machines. At worst, it leads researchers to make plainly false claims about their creations; for example, Yamamoto described his robot vacuum cleaner Sozzy as 'friendly' [153] and Hogg, Martin, and Resnick said that Frantic, their Braitenberg-like creature made of Lego bricks, 'does nothing but think' [84].

In a classic 1976 paper entitled 'Artificial Intelligence Meets Natural Stupidity', Drew McDermott advised scientists to use 'colourless' or 'sanitized' technical descriptions of their machines in place of unreflective and misleading psychological expressions [98, p. 4]. (McDermott's target was 'wishful mnemonics' [98, p. 4] but anthropomorphizing in AI goes far beyond such shorthand.) Several researchers in social robotics can be seen as in effect attempting to follow McDermott's advice with respect to their anthropomorphic robots. These researchers refrain from saying that their 'expressive' robots have emotions, and instead say that they have emotional behaviour. Kismet, Cynthia Breazeal's famous (now retired) interactive 'expressive' robot was said (without scare-quotes) to have a 'smile on [its] face', a 'sorrowful expression', a 'fearful expression', a 'happy and interested expression', a 'contented smile', a 'big grin', and a 'frown' ([31, p. 584–588], [30,35]). However, this vocabulary is not sufficiently sanitized: for example, to say that a machine smiles is to say that the machine has an intent, namely to communicate, and an inner state, typically happiness.[4] Here the forensic problem of anthropomorphism reemerges. We need a test for expressiveness, as much as for intelligence, that is not undermined by our tendency to anthropomorphize.

5.2 Uncanny Valley

Anthropomorphism not only affects how people behave towards robots, but also whether they will accept them in natural human environments. The relation between the physical appearance of robots and their acceptance has recently received major interest in the field of HRI. Despite this, there are still many unanswered questions and most efforts are devoted to the uncanny valley theory [99]. This theory proposes a non-linear relationship between a robot's degree of anthropomorphism and its likeability. With increased human-likeness a robot's likeability also increases; yet when a robot closely resembles a human, but is not identical, this produces a strong negative emotional reaction. Once a robot's appearance is indistinguishable from that of a human, the robot is liked as much as a human being [99].

It has been suggested that neurological changes are responsible for the uncanny valley phenomenon [124]. Furthermore, it is the perceived higher ability of anthropomorphic robots to have experience that makes people particularly uncomfortable with human-like technology [73]. However, in spite of its popularity the empirical proof of the uncanny valley theory is relatively sparse. Some studies did not find evidence supporting this hypothesis [19], while others suggest that the relation between likeability and appearance might have a different shape, one that resembles more a cliff than a valley [18]. We believe that future research should address three key issues: defining terminology, finding entities that lie between the deepest point of the uncanny valley and the human level, and investigating the uncanny valley in studies that involve actual HRI.

Up to now multiple terms have been used in place of the Japanese term used by Mori to describe the uncanny valley. This reduces the comparability of the studies. Moreover, other researchers point out that even the human-likeness axis of the graph is not well-defined [38]. Efforts are spent on trying to find a term that would fit the hypothesized shape of the valley rather than on creating a hypothesis that fits the data. It is also possible that the term used by Mori might not be the most appropriate one and that the problem does not lie only in the translation.

The uncanny valley hypothesis suggests that when a robot crosses the deepest point of the uncanny valley its likeability will suddenly increase. However, to date, no entity has

---

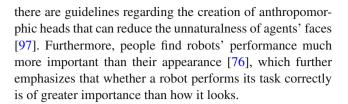[2] Daniel Dennett uses the notion of 'innocent' anthropomorphizing in [46].

[3] Turing distinguished between communication by 'pain' and 'pleasure' inputs and 'unemotional' communication that by means of 'sense stimuli' [141]; for analysis see [114].

[4] This is not to suggest that what makes a 'smile' into a smile is some feeling—an inner state—in the robot. See [111,113].

been shown to exist that is similar enough to a human for it to fit this description. We propose that work on the opposite process to anthropomorphism could fill that niche. It has been suggested that dehumanization, which is the deprivation of human qualities in real human beings, is such a process [78]. Work on dehumanization shows which characteristics are perceived as critical for the perception of others as human. Their elimination leads to people being treated as if they were not fully human. The studies of dehumanization show that there are two distinct senses of humanness—uniquely human and human nature. Uniquely human characteristics distinguish humans from other species and reflect attributes such as intelligence, intentionality, or secondary emotions. On the other hand, people deprived of human nature are perceived as automata and lacking in primary emotions, sociability, or warmth. These two dimensions map well onto the concept proposed for the dimensionality of mind-attribution that was found to involve agency and experience [72]. Therefore, we could explore the uncanny valley, not by trying to reach the human level starting from a machine, but rather by using humans that are perceived as lacking some human qualities. There is some empirical evidence that anthropomorphism is also a multi-dimensional phenomenon [156].

In addition, all previous studies of the uncanny valley hypothesis have used either static images or videos of robots. The question remains how well these findings can be generalized to actual HRI. It is possible that the uncanny valley would have no effect on HRI or that it would be limited to the very first few seconds of interaction. The studies of the uncanny valley phenomenon in computer graphics indicate that this phenomenon might be related to the exposure to a specific agent [47]. Increased familiarity with an agent could be related with decreased uncanniness felt as a result of its appearance. The physical appearance of a robot is not the most important factor in anthropomorphism [90]. Furthermore, the perception of human-likeness changes during interaction [68]. It is possible that the uncanny valley might lead to people being less willing to engage in interaction. However, we believe that more effort should be put into interaction design rather than physical-appearance design, since the relationship between the former and the uncanny valley needs further empirical research.

6 Overcoming the Problems of Anthropomorphic Technology

Even if we accept the uncanny valley as it was proposed by Mori, there are certain reasons why the consequences for acceptance of anthropomorphic robots are not as profound as the theory indicated. In non-laboratory conditions people rarely reported an eerie feeling when interacting with a geminoid [21]. Furthermore, at least for computer graphics

there are guidelines regarding the creation of anthropomorphic heads that can reduce the unnaturalness of agents' faces [97]. Furthermore, people find robots' performance much more important than their appearance [76], which further emphasizes that whether a robot performs its task correctly is of greater importance than how it looks.

6.1 Facilitating positive attitudes toward eerie machines

If the uncanny valley has a lasting effect on HRI, it is beneficial to consider how the acceptance of eerie machines could be facilitated. Previous work in HRI shows that people can perceive robots as either ingroup or outgroup members [57] and even apply racial prejudice towards them [58]. Therefore, the theoretical foundations for the integration of highly human-like robots could build on the extensive research examining how to facilitate positive intergroup relations between humans belonging to differing nationalities, ethnicities, sexual, or religious groups. The topic of intergroup relations has been heavily investigated by social psychologists worldwide since World War II. While some of this early work was interested in understanding psychological factors that led to the events of the Holocaust, Gordon Allport's seminal work on the nature of prejudice [5] provided the field with a larger platform to examine the basic psychological factors underlying stereotyping, prejudice, and discrimination.

From several decades of research on the topic, the field has not only shed light on the varied ways in which intergroup bias manifests itself in everyday life [45,74], but it also helps us better understand the economic, motivational, cognitive, evolutionary, and ideological factors that drive intergroup bias and conflict between social groups [83]. In addition, the field has also identified several social psychological approaches and strategies that can be used to reduce prejudice, stereotyping, and discrimination toward outgroups (i.e. groups to which we do not belong). These strategies range from interventions that promote positive feelings and behaviour toward outgroups through media messages [50,105,152], recategorization of outgroups into a common superordinate group [50,51], valuing diversity and what each subgroup can contribute to the greater good [145,146,154], promoting positive contact with members of the outgroup [50,106,107], among others.

In the context of HRI, these social psychological strategies may be used to promote positive HRI and favorable social attitudes toward robots. For example, from over fifty years of empirical research on intergroup contact, there is strong empirical evidence that positive contact with an outgroup can reduce prejudice or negative feelings toward the outgroup [106,107]. Such positive contact between two groups has been shown to be particularly beneficial when four conditions are met: (a) within a given situation, the perceived

status of the two groups must be equal; (b) they must have common goals; (c) cooperation between the two groups must occur; and (d) positive contact between groups must be perceived as sanctioned by authority. Such intergroup contact may reduce outgroup prejudice for many different reasons [106,107], one reason being that positive contact allows one to learn more about the outgroup. In the context of HRI, one may, therefore, expect that negative attitudes towards human-like robots may stem from their perceived unfamiliarity and unpredictability. Although they look human-like, people cannot be sure whether machines will behave like a human being. Increased familiarity with such technology might thereby lead to decreased uncertainty regarding its actions and in turn reduce negative feelings toward them. Empirical research is needed in order to establish whether intergroup contact can facilitate greater acceptance of anthropomorphic robots. More broadly, such social psychological research may offer insight into understanding when and why people may feel unfavorably toward robots, while offering practical strategies that can be considered in HRI as a means to promote greater social acceptance of robots in our increasingly technological world.

6.2 Limiting the Risks Associated with Anthropomorphic Technology

Of the criticism that anthropomorphic robots (in fiction at least) encourage the general public to think that AI has progressed further than in actuality, we can simply say that this may underestimate the public's good sense. The objection, against those researchers aiming to build a child-machine, that human-like AI is a mistaken and unproductive goal can also be answered. For example, the real target of this complaint may be, not human-level or human-like AI as such, but rather symbolic AI as a means of attaining human-level AI (see [110]). Behaviour-based approaches may escape this complaint. Moreover, the assumption that there is such a thing as generic intelligence, and this is the proper subject of study for researchers in computational intelligence, begs an important question. Perhaps our concept of intelligence just is drawn from the paradigm examples of thinking things—human beings.

This leaves the forensic problem of anthropomorphism. In general, AI requires a distinction between a mere and a thinking machine, and this distinction must be proof against the human tendency to anthropomorphize. This is exactly what Turing's imitation game provides (see [110,112]). The game disincentivizes anthropomorphism: an observer (i.e. interrogator) who anthropomorphizes a contestant increases the chances of making the embarrassing mistake of misidentifying a computer as a human being. The behaviour of interrogators (in early Loebner Prize Contests where a series of machine and human contestants were interviewed individu-

ally) shows that observers go out of their way to avoid this awkward error, to the extent that they misidentify human beings as computers. In addition, the imitation game controls for the effect of the tendency to anthropomorphize; in simultaneous interviews with a machine and a human contestant, an observer's propensity to anthropomorphize (which we can assume to be present equally in both interviews) cannot advantage one contestant over the other. Turing's test is proofed against the human tendency to anthropomorphize machines.

But how is anthropomorphism-proofing to be applied to judgements of intelligence or affect in anthropomorphic robots? Much of the engineering of these robots is to make them visibly indistinguishable from a human being. An open imitation game where both contestants—a hyper-realistic anthropomorphic robot and an actual human being—are seen by the interrogator would provide a disincentive to anthropomorphizing and a control on the tendency to anthropomorphize. However, interrogators in this game might well focus on characteristics of the contestants that Turing labeled 'irrelevant' disabilities—qualities immaterial to the question whether a machine can think, such as a failure 'to shine in beauty competitions' [139, p. 442]. An interrogator might, for example, concentrate on the functioning of a contestant's facial muscles or the appearance of the skin. This open game, although anthropomorphism-proofed, would fail as a test of intelligence or affect in machines. On the other hand, in the standard game where both the robot and the human contestants are hidden, much of the robot's engineering would now be irrelevant to its success or failure in the game—for example, David Hanson's robot Einstein's 'eyebrows' [103] surely do not contribute to a capacity for cognition or affect. This is why Turing said that there is 'little point in trying to make a "thinking machine" more human by dressing it up in ... artificial flesh' [139, p. 442] and hoped that 'no great efforts will be put into making machines with the most distinctively human, but non-intellectual characteristics such as the shape of the human body' [143, p. 486].

In sum, AI requires some means of anthropomorphism-proofing judgements of intelligence or affect in anthropomorphic robots—otherwise it lacks a distinction between justified and unjustified anthropomorphizing. An open Turing test will test for the wrong things. A standard Turing test will suffice, but seems to be inconsistent with the growing trend for hyper-realistic and eerie robots.

7 Conclusions

In this paper we have discussed the widespread tendency for people to anthropomorphise their surroundings and in particular how this affects HRI. Our understanding of its impact on HRI is still in its infancy. However, there is no doubt that it

creates new opportunities and poses problems that can have profound consequences for the field of HRI and acceptance of the robotic technology.

Anthropomorphism is not only limited to the appearance of a robot, but the design of a robotic platform must also consider a robot's interaction with humans as an important factor. Accordance between these factors is necessary for a robot to maintain its human-like impression. A well designed system can facilitate interaction, but it must match the specific task given to a robot. For people it is more important that a robot can do its job accurately rather than how it looks. However, we have presented multiple examples where anthropomorphic form in appearance and behavior can help a robot to perform its tasks successfully by eliciting desired behaviours from human interaction partners.

On the other hand, development of anthropomorphic robots comes at certain costs. People expect them to adhere to human norms and have much higher expectations regarding their capabilities compared to robots with machine-like appearance. The uncanny valley hypothesis suggests that there is repulsion toward highly human-like machines that are still distinguishable from humans. However, in this paper we have shown the main shortcoming of the previous work that might hamper the suitability of this theory in HRI. Future research should focus on investigating of this phenomenon in real HRI rather than by using images or videos. Moreover, work on the opposite process, dehumanization, can help us to understand the relationship between acceptance and anthropomorphism better. In addition, in order to facilitate the integration of human-like robots we propose to employ strategies from the area of intergroup relations that are being used to facilitate positive relations between human subgroups.

We have also shown that the phenomenon of anthropomorphic robots generates challenging philosophical and psychological questions. In order for the field of AI to progress further it is necessary to acknowledge them. These challenges can not only affect how the general public perceives current and future directions of research and anthropomorphic and intelligent systems, but also might determine how far the field can go. It remains to be seen whether the field can successfully address these problems.

References

1. Aaker J (1997) Dimensions of brand personality. J Mark Res 34:347–356
2. Abend L (2008) In spain, human rights for apes. TIMEcom http://www.time.com/time/world/article/08599(1824206)
3. Adolphs R (2005) Could a robot have emotions? Theoretical perspectives from social cognitive neuroscience. In: Arbib M, Fellous JM (eds) Who needs emotions: the brain meets the robot. Oxford University Press, Oxford, pp 9–28
4. Aggarwal P, McGill A (2007) Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. J Consum Res 34(4):468–479
5. Allport GW (1954) The nature of prejudice. Addison-Wesley, Reading
6. Arbib MA, Fellous JM (2004) Emotions: from brain to robot. Trends Cogn Sci 8(12):554–561
7. Austermann A, Yamada S, Funakoshi K, Nakano M (2010) How do users interact with a pet-robot and a humanoid. In: Conference on human factors in computing systems—Proceedings, Atlanta, pp 3727–3732
8. Baddoura R, Venture G, Matsukata R (2012) The familiar as a key-concept in regulating the social and affective dimensions of HRI. In: IEEE-RAS international conference on humanoid robots, Osaka, pp 234–241
9. Bae J, Kim M (2011) Selective visual attention occurred in change detection derived by animacy of robot's appearance. In: Proceedings of the 2011 international conference on collaboration technologies and systems, CTS 2011, pp 190–193
10. Barrett J (2004) Why would anyone believe in god?. AltaMira Press, Lanham
11. Barrett J (2007) Cognitive science of religion: what is it and why is it? Relig Compass 1(6):768–786
12. Barrett JL (2000) Exploring the natural foundations of religion. Trends Cogn Sci 4(1):29–34
13. Bartneck C (2008) Who like androids more: Japanese or US americans? Proceedings of the 17th IEEE international symposium on robot and human interactive communication, RO-MAN, Munich, pp 553–557
14. Bartneck C (2013) Robots in the theatre and the media. In: Design & semantics of form & movement (DeSForM2013), Philips, pp 64–70
15. Bartneck C, Hue J (2008) Exploring the abuse of robots. Interact Stud 9(3):415–433
16. Bartneck C, Rosalia C, Menges R, Deckers I (2005) Robot abuse: limitation of the media equation. In: Proceedings of the interact 2005 workshop on agent abuse, Rome
17. Bartneck C, Reichenbach J, Carpenter J (2006) Use of praise and punishment in human–robot collaborative teams. In: RO-MAN 2006: the 15th IEEE international symposium on robot and human interactive communication (IEEE Cat No. 06TH8907), Piscataway
18. Bartneck C, Kanda T, Ishiguro H, Hagita N (2007) Is the uncanny valley an uncanny cliff? In: Proceedings of the IEEE international workshop on robot and human interactive communication, Jeju, pp 368–373
19. Bartneck C, Kanda T, Ishiguro H, Hagita N (2009) My robotic doppelganger - a critical look at the uncanny valley theory. In: 18th IEEE international symposium on robot and human interactive communication, RO-MAN2009. IEEE, pp 269–276
20. Bartneck C, Bleeker T, Bun J, Fens P, Riet L (2010) The influence of robot anthropomorphism on the feelings of embarrassment when interacting with robots. Paladyn 1:1–7
21. Becker-Asano C, Ogawa K, Nishio S, Ishiguro H (2010) Exploring the uncanny valley with geminoid HI-1 in a real-world application. In: Proceedings of the IADIS international conference interfaces and human computer interaction 2010, IHCI. Proceedings of the IADIS international conference game and entertainment technologies 2010, Part of the MCCSIS 2010, Freiburg, pp 121–128
22. Bering J (2005) Origins of the social mind: evolutionary psychology and child development. The Guildford Press, New York
23. Bering J (2010) The god instinct. Nicholas Brealey, London
24. Bering JM (2006) The folk psychology of souls. Behav Brain Sci 29(05):453–462

25. Bethel CL, Salomon K, Murphy RR (2009) Preliminary results: humans find emotive non-anthropomorphic robots more calming. In: Proceedings of the 4th ACM/IEEE international conference on human–robot interaction, HRI'09, pp 291–292

26. Bloom P (2005) Descartes' baby: how the science of child development explains what makes us human. Basic Books, New York

27. Boyer P (2001) Religion explained. Basic Books, New York

28. Boyer P (2003) Religious thought and behaviour as by-products of brain function. Trends Cogn Sci 7(3):119–124

29. Breazeal C (1998) Early experiments using motivations to regulate human–robot interaction. In: AAAI fall symposium on emotional and intelligent: the tangled knot of cognition. Technical report FS-98-03, pp 31–36

30. Breazeal C (2000) Sociable machines: expressive social exchange between humans and robots. Dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology

31. Breazeal C (2001) Affective interaction between humans and robots. In: Kelemen J, Sosik P (eds) ECAL 2001. LNAI vol 2159. Springer-Verlag, Berlin, pp 582–591

32. Breazeal C (2006) Human–robot partnership. IEEE Intell Syst 21(4):79–81

33. Breazeal C, Fitzpatrick P (2000) That certain look: social amplification of animate vision. In: Proceedings of the AAAI fall symposium on society of intelligence agents the human in the loop

34. Breazeal C, Scassellati B (2000) Infant-like social interactions between a robot and a human caregiver. Adapt Behav 8(1):49–74

35. Breazeal C, Scassellati B (2001) Challenges in building robots that imitate. In: Dautenhahn K, Nehaniv CL (eds) Imitation in animals and artifacts. MIT Press, Cambridge

36. Breazeal C, Buchsbaum D, Gray J, Gatenby D, Blumberg B (2005) Learning from and about others: towards using imitation to bootstrap the social understanding of others by robots. Artif Life 11(1–2):31–62

37. Calverley DJ (2006) Android science and animal rights, does an analogy exist? Connect Sci 18(4):403–417

38. Cheetham M, Suter P, Jancke L (2011) The human likeness dimension of the "Uncanny valley hypothesis": behavioral and functional MRI findings. Front Hum Neurosci 5:126

39. Chew S, Tay W, Smit D, Bartneck C (2010) Do social robots walk or roll? In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and Lecture notes in bioinformatics), LNAI, vol 6414. Springer, Singapore, pp 355–361

40. Chin M, Sims V, Clark B, Lopez G (2004) Measuring individual differences in anthropomorphism toward machines and animals. In: Proceedings of the human factors and ergonomics society annual meeting, vol 48. SAGE Publications, Thousand Oaks, pp 1252–1255

41. Choi J, Kim M (2009) The usage and evaluation of anthropomorphic form in robot design. In: Undisciplined! Design research society conference 2008, Sheffield Hallam University, Sheffield, 16–19 July 2008

42. Cohen PR (2005) If not turing's test, then what? AI Mag 26(4):61

43. Cooney M, Zanlungo F, Nishio S, Ishiguro H (2012) Designing a flying humanoid robot (FHR): effects of flight on interactive communication. In: Proceedings of the IEEE international workshop on robot and human interactive communication, Paris, pp 364–371

44. Darwin C (1872) 1998 The expression of the emotions in man and animals. Oxford University Press, New York

45. Dasgupta N (2004) Implicit in group favoritism, outgroup favoritism, and their behavioral manifestations. Soc Justice Res 17(2):143–169

46. Dennett DC, Dretske F, Shurville S, Clark A, Aleksander I, Cornwell J (1994) The practical requirements for making a conscious robot. Philos Trans R Soc Lond Ser A 349(1689):133–146

47. Dill V, Flach LM, Hocevar R, Lykawka C, Musse SR, Pinho MS (2012) Evaluation of the uncanny valley in CG characters. In: 12th international conference on intelligent virtual agents, IVA 2012, September 12, 2012–September 14, 2012. Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and Lecture notes in bioinformatics) LNAI, vol 7502, Springer Verlag, Berlin, pp 511–513

48. DiSalvo C, Gemperle F (2003) From seduction to fulfillment: the use of anthropomorphic form in design. In: Proceedings of the 2003 international conference on designing pleasurable products and interfaces, DPPI '03, ACM, New York, pp 67–72

49. DiSalvo CF, Gemperle F, Forlizzi J, Kiesler S (2002) All robots are not created equal: the design and perception of humanoid robot heads. In: Proceedings of the conference on designing interactive systems: processes, practices, methods, and techniques. DIS, London, pp 321–326

50. Dovidio J, Gaertner S (1999) Reducing prejudice: combating intergroup biases. Curr Dir Psychol Sci 8(4):101–105

51. Dovidio J, Gaertner S, Saguy T (2009) Commonality and the complexity of "we" social attitudes and social change. Pers Soc Psychol Rev 13(1):3–20

52. Duffy BR (2003) Anthropomorphism and the social robot. Robot Auton Syst 42(3—-4):177–190

53. Epley N, Waytz A, Cacioppo JT (2007) On seeing human: a three-factor theory of anthropomorphism. Psychol Rev 114(4): 864–886

54. Epley N, Akalis S, Waytz A, Cacioppo J (2008) Creating social connection through inferential reproduction: loneliness and perceived agency in gadgets, gods, and hreyhounds: Research article. Psychol Sci 19(2):114–120

55. Evers V, Maldonado HC, Brodecki TL, Hinds PJ (2008) Relational vs. group self-construal: untangling the role of national culture in HRI. In: HRI 2008: Proceedings of the 3rd ACM/IEEE international conference on human–robot interaction: living with robots, Amsterdam, pp 255–262

56. Eyssel F, Kuchenbrandt D (2011) Manipulating anthropomorphic inferences about NAO: the role of situational and dispositional aspects of effectance motivation. In: Proceedings of the IEEE international workshop on robot and human interactive communication, pp 467–472

57. Eyssel F, Kuchenbrandt D (2012) Social categorization of social robots: anthropomorphism as a function of robot group membership. Br J Soc Psychol 51(4):724–731

58. Eyssel F, Loughnan S (2013) "It don't matter if you're black or white"? Effects of robot appearance and user prejudice on evaluations of a newly developed robot companion. In: 5th international conference on social robotics, ICSR 2013, October 27, 2013–October 29, 2013. Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and Lecture notes in bioinformatics). LNAI, vol 8239. Springer Verlag, Berlin, pp 422–431

59. Eyssel F, Hegel F, Horstmann G, Wagner C (2010) Anthropomorphic inferences from emotional nonverbal cues: A case study. In: Proceedings of the IEEE international workshop on robot and human interactive communication. Viareggio, pp 646–651

60. Eyssel F, Kuchenbrandt D, Bobinger S (2011) Effects of anticipated human–robot interaction and predictability of robot behavior on perceptions of anthropomorphism. In: HRI 2011: Proceedings of the 6th ACM/IEEE international conference on human–robot interaction. Lausanne, pp 61–67

61. Eyssel F, Kuchenbrandt D, Bobinger S, De Ruiter L, 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. In: HRI'12: Proceedings of the 7th annual ACM/IEEE international conference on human–robot interaction, pp 125–126

62. Fasola J, Matarić MJ (2012) Using socially assistive human–robot interaction to motivate physical exercise for older adults. Proc IEEE 100:2512–2526

63. Feil-Seifer D, Matarić MJ (2011) Automated detection and classification of positive vs. negative robot interactions with children with autism using distance-based features. In: HRI 2011: Proceedings of the 6th ACM/IEEE international conference on human–robot interaction, pp 323–330

64. Fischer K, Lohan KS, Foth K (2012) Levels of embodiment: Linguistic analyses of factors influencing HRI. In: HRI'12: Proceedings of the 7th annual ACM/IEEE international conference on human–robot interaction, pp 463–470

65. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. In: IROS 2002, September 30, 2002–September 30, 2002. Robotics and autonomous systems, vol 42. Elsevier, pp 143–166

66. Ford K, Hayes P (1998) On computational wings: rethinking the goals of artificial intelligence-the gold standard of traditional artificial intelligence-passing the so-called turing test and thereby appearing to be. Sci Am Presents 9(4):79

67. Forster F, Weiss A, Tscheligi M (2011) Anthropomorphic design for an interactive urban robot—the right design approach? In: HRI 2011: Proceedings of the 6th ACM/IEEE international conference on human–robot interaction, Lausanne, pp 137–138

68. Fussell SR, Kiesler S, Setlock LD, Yew V (2008) How people anthropomorphize robots. In: HRI 2008: Proceedings of the 3rd ACM/IEEE international conference on human–robot interaction: living with robots, Amsterdam, pp 145–152

69. Giullian N, Ricks D, Atherton A, Colton M, Goodrich M, Brinton B (2010) Detailed requirements for robots in autism therapy. In: Conference Proceedings: IEEE international conference on systems, man and cybernetics, pp 2595–2602

70. Goetz J, Kiesler S, Powers A (2003) Matching robot appearance and behavior to tasks to improve human–robot cooperation. In: Proceedings of the 12th IEEE international workshop on robot and human interactive communication (ROMAN 2003), 2003, pp 55–60

71. Gold K, Scassellati B (2007) A bayesian robot that distinguishes "self" from "other". In: Proceedings of the 29th annual meeting of the cognitive science society

72. Gray HM, Gray K, Wegner DM (2007) Dimensions of mind perception. Science 315(5812):619

73. Gray K, Wegner D (2012) Feeling robots and human zombies: mind perception and the uncanny valley. Cognition 125(1):125–130

74. Greenwald A, Banaji M (1995) Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychol Rev 102(1):4–27

75. Guthrie S (1995) Faces in the clouds: a new theory of religion. Oxford University Press, Oxford

76. Hancock PA, Billings DR, Schaefer KE, Chen JYC (2011) A meta-analysis of factors affecting trust in human–robot interaction. Hum Factors 53(5):517–527

77. Hard R (2004) The routledge handbook of greek mythology: Based on HJ Rose's "Handbook of Greek Mythology". Psychology Press, New York

78. Haslam N (2006) Dehumanization: an integrative review. Pers Soc Psychol Rev 10(3):252–264

79. Hayes P, Ford K (1995) Turing test considered harmful. In: IJCAI (1), pp 972–977.

80. Hegel F, Krach S, Kircher T, Wrede B, Sagerer G (2008) Understanding social robots: a user study on anthropomorphism. In: Proceedings of the 17th IEEE international symposium on robot and human interactive communication, RO-MAN, pp 574–579

81. Hegel F, Gieselmann S, Peters A, Holthaus P, Wrede B (2011) Towards a typology of meaningful signals and cues in social robotics. In: Proceedings of the IEEE international workshop on robot and human interactive communication, pp 72–78

82. Heider F, Simmel M (1944) An experimental study of apparent behavior. Am J Psychol 57(2):243–259

83. Hewstone M, Rubin M, Willis H (2002) Intergroup bias. Annu Rev Psychol 53:575–604

84. Hogg DW, Martin F, Resnick M (1991) Braitenberg creatures. Epistemology and Learning Group, MIT Media Laboratory, Cambridge

85. Ishiguro H (2006) Android science: conscious and subconscious recognition. Connect Sci 18(4):319–332

86. Kahn P, Ishiguro H, Friedman B, Kanda T (2006) What is a human? Toward psychological benchmarks in the field of human–robot interaction. In: The 15th IEEE international symposium on robot and human interactive communication, 2006. ROMAN 2006, pp 364–371

87. Kahn PH Jr, Kanda T, Ishiguro H, Gill BT, Ruckert JH, Shen S, Gary HE, Reichert AL, Freier NG, Severson RL (2012) Do people hold a humanoid robot morally accountable for the harm it causes? In: HRI'12: Proceedings of the 7th annual ACM/IEEE international conference on human–robot interaction, Boston, pp 33–40

88. Kamide H, Kawabe K, Shigemi S, Arai T (2013) Development of a psychological scale for general impressions of humanoid. Adv Robot 27(1):3–17

89. Kanda T, Miyashita T, Osada T, Haikawa Y, Ishiguro H (2005) Analysis of humanoid appearances in human–robot interaction. In: 2005 IEEE/RSJ international conference on intelligent robots and systems. IROS, Edmonton, pp 62–69

90. Kiesler S, Goetz J (2002) Mental models and cooperation with robotic assistants. In: Proceedings of the conference on human factors in computing systems, pp 576–577

91. Kiesler S, Hinds P (2004) Introduction to this special issue on human–robot interaction. Hum-Comput Interact 19(1):1–8

92. Kiesler S, Powers A, Fussell SR, Torrey C (2008) Anthropomorphic interactions with a robot and robot-like agent. Soc Cogn 26(2):169–181

93. Kuchenbrandt D, Eyssel F, Bobinger S, Neufeld M (2013) When a robot's group membership matters. Int J Soc Robot 5(3):409–417

94. Lee Sl, Lau I, Kiesler S, Chiu CY (2005) Human mental models of humanoid robots. In: Proceedings of the 2005 IEEE international conference on robotics and automation, 2005. ICRA 2005, pp 2767–2772

95. Levin D, Killingsworth S, Saylor M, Gordon S, Kawamura K (2013) Tests of concepts about different kinds of minds: predictions about the behavior of computers, robots, and people. Hum Comput Interact 28(2):161–191

96. Lohse M (2011) Bridging the gap between users' expectations and system evaluations. In: Proceedings of the IEEE international workshop on robot and human interactive communication, pp 485–490

97. MacDorman KF, Green RD, Ho CC, Koch CT (2009) Too real for comfort? Uncanny responses to computer generated faces. Comput Hum Behav 25(3):695–710

98. McDermott D (1976) Artificial intelligence meets natural stupidity. ACM SIGART Bull 57:4–9

99. Mori M (1970) The uncanny valley. Energy 7(4):33–35

100. Mutlu B, Yamaoka F, Kanda T, Ishiguro H, Hagita N (2009) Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In: Proceedings of the 4th ACM/IEEE international conference on human robot interaction, HRI '09. ACM, New York, pp 69–76

101. Oestreicher L, Eklundh KS (2006) User expectations on human-robot co-operation. In: Robot and human interactive communica-

tion, 2006. The 15th IEEE international symposium on ROMAN 2006. IEEE, pp 91–96

102. Ogawa K, Bartneck C, Sakamoto D, Kanda T, Ono T, Ishiguro H (2009) Can an android persuade you? In: RO-MAN 2009: the 18th IEEE international symposium on robot and human interactive communication, Piscataway, pp 516–521

103. Oh JH, Hanson D, Kim WS, Han IY, Kim JY, Park IW (2006) Design of android type humanoid robot albert HUBO. In: 2006 IEEE/RSJ international conference on intelligent robots and systems, pp 1428–1433

104. Oztop E, Chaminade T, Franklin D (2004) Human-humanoid interaction: is a humanoid robot perceived as a human? In: 2004 4th IEEE/RAS international conference on humanoid robots, vol 2, pp 830–841

105. Paluck E (2009) Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda. J Pers Soc Psychol 96(3):574–587

106. Pettigrew T (1998) Intergroup contact theory. Annu Rev Psychol 49:65–85

107. Pettigrew T, Tropp L (2006) A meta-analytic test of intergroup contact theory. J Pers Soc Psychol 90(5):751–783

108. Pollack JB (2006) Mindless intelligence. IEEE Intell Syst 21(3):50–56

109. Powers A, Kramer ADI, Lim S, Kuo J, Lee SL, Kiesler S (2005) Eliciting information from people with a gendered humanoid robot. In: Proceedings of the IEEE international workshop on robot and human interactive communication, vol 2005. Nashville, pp 158–163

110. Proudfoot D (2011) Anthropomorphism and AI: turing's much misunderstood imitation game. Artif Intell 175(5):950–957

111. Proudfoot D (2013a) Can a robot smile? Wittgenstein on facial expression. In: Racine TP, Slaney KL (eds) A Wittgensteinian perspective on the use of conceptual analysis in psychology. Palgrave Macmillan, Basingstoke, pp 172–194

112. Proudfoot D (2013b) Rethinking turing's test. J Philos 110(7):391–411

113. Proudfoot D (2015) Turing's child-machines. In: Bowen J, Copeland J, Sprevak M, Wilson R (eds) The turing guide: life, work, legacy. Oxford University Press, Oxford

114. Proudfoot D (2014) Turing's three senses of "Emotional". Int J Synth Emot 5(2):7–20

115. Pyysiäinen I (2004) Religion is neither costly nor beneficial. Behav Brain Sci 27(06):746–746

116. Reeves B, Nass C (1996) The media equation. Cambridge University Press, New York

117. Rehm M, Krogsager A (2013) Negative affect in human robot interaction—impoliteness in unexpected encounters with robots. In: 2013 IEEE RO-MAN, pp 45–50

118. Reichenbach J, Bartneck C, Carpenter J (2006) Well done, robot! - The importance of praise and presence in human–robot collaboration. In: Proceedings of the IEEE international workshop on robot and human interactive communication, Hatfield, pp 86–90

119. Riek LD, Rabinowitch TC, Chakrabarti B, Robinson P (2008) How anthropomorphism affects empathy toward robots. In: Proceedings of the 4th ACM/IEEE international conference on human–robot interaction, HRI'09, San Diego, pp 245–246

120. Riether N, Hegel F, Wrede B, Horstmann G (2012) Social facilitation with social robots? In: Proceedings of the seventh annual ACM/IEEE international conference on human–robot interaction, ACM, Boston, HRI '12, pp 41–48

121. Saerbeck M, Schut T, Bartneck C, Janse MD (2010) Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In: Proceedings of the conference on human factors in computing systems, vol 3, pp 1613–1622

122. Salem M, Eyssel F, Rohlfing K, Kopp S, Joublin F (2011) Effects of gesture on the perception of psychological anthropomorphism: a case study with a humanoid robot, Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and Lecture notes in bioinformatics). LNAI, vol 7072. Springer, Berlin

123. Saver JL, Rabin J (1997) The neural substrates of religious experience. J Neuropsychiatry Clin Neurosci 9(3):498–510

124. Saygin AP, Chaminade T, Ishiguro H, Driver J, Frith C (2012) The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. Soc Cogn Affect Neurosci 7(4):413–422

125. Scassellati B (2000) How robotics and developmental psychology complement each other. In: NSF/DARPA workshop on development and learning

126. Scassellati B (2002) Theory of mind for a humanoid robot. Auton Robots 12(1):13–24

127. Scassellati B (2007) How social robots will help us to diagnose, treat, and understand autism. Springer Tracts in Advanced Robotics, vol 28, pp 552–563

128. Scassellati B, Crick C, Gold K, Kim E, Shic F, Sun G (2006) Social development [robots]. Comput Intell Mag IEEE 1(3):41–47

129. Schmitz M (2011) Concepts for life-like interactive objects. In: Proceedings of the fifth international conference on tangible, embedded, and embodied interaction, TEI '11. ACM, New York, pp 157–164

130. Shic F, Scassellati B (2007) Pitfalls in the modeling of developmental systems. Int J Hum Robot 4(2):435–454

131. Short E, Hart J, Vu M, Scassellati B (2010) No fair!! An interaction with a cheating robot. In: 5th ACM/IEEE international conference on human–robot interaction, HRI 2010, Osaka, pp 219–226

132. Sims VK, Chin MG, Lum HC, Upham-Ellis L, Ballion T, Lagattuta NC (2009) Robots' auditory cues are subject to anthropomorphism. In: Proceedings of the human factors and ergonomics society, vol 3, pp 1418–1421

133. Spexard T, Haasch A, Fritsch J, Sagerer G (2006) Human-like person tracking with an anthropomorphic robot. In: Proceedings of the IEEE international conference on robotics and automation, vol 2006. Orlando, pp 1286–1292

134. Syrdal DS, Dautenhahn K, Woods SN, Walters ML, Koay KL (2007) Looking good? appearance preferences and robot personality inferences at zero acquaintance. In: AAAI spring symposium—Technical report, vol SS-07-07. Stanford, pp 86–92

135. Syrdal DS, Dautenhahn K, Walters ML, Koay KL (2008) Sharing spaces with robots in a home scenario—anthropomorphic attributions and their effect on proxemic expectations and evaluations in a live HRI trial. In: AAAI fall symposium—Technical report, vol FS-08-02, Arlington, pp 116–123

136. Tapus A, Matari MJ, Scassellati B (2007) Socially assistive robotics [grand challenges of robotics]. IEEE Robotics Autom Mag 14(1):35–42

137. Torta E, Van Dijk E, Ruijten PAM, Cuijpers RH (2013) The ultimatum game as measurement tool for anthropomorphism in human–robot interaction. In: 5th international conference on social robotics, ICSR 2013, October 27, 2013–October 29, 2013. Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and Lecture notes in bioinformatics). LNAI, vol 8239. Springer Verlag, pp 209–217

138. Trimble M, Freeman A (2006) An investigation of religiosity and the Gastaut-Geschwind syndrome in patients with temporal lobe epilepsy. Epilepsy Behav 9(3):407–414

139. Turing A (1950) Computing machinery and intelligence. Mind 59(236):433–460

140. Turing A (2004a) Can digital computers think? In: Copeland BJ (ed) The essential turing. Oxford University Press, Oxford

141. Turing A (2004b) Intelligent machinery. In: Copeland BJ (ed) The essential turing. Oxford University Press, Oxford

142. Turing A (2004c) Lecture on the automatic computing engine. In: Copeland BJ (ed) The essential turing. Oxford University Press, Oxford

143. Turing A, Braithwaite R, Jefferson G, Newman M (2004) Can automatic calculating machines be said to think? In: Copeland J (ed) The essential turing. Oxford University Press, Oxford

144. Turkle S (2010) In good company? On the threshold of robotic companions. In: Wilks Y (ed) Close engagements with artifical companions: key social, psychological, ethical and design issues. John Benjamins Publishing Company, Amsterdam/ Philadelphia, pp 3–10

145. Verkuyten M (2006) Multicultural recognition and ethnic minority rights: a social identity perspective. Eur Rev Soc Psychol 17(1):148–184

146. Vorauer J, Gagnon A, Sasaki S (2009) Salient intergroup ideology and intergroup interaction. Psychol Sci 20(7):838–845

147. Wade E, Parnandi AR, Mataric MJ (2011) Using socially assistive robotics to augment motor task performance in individuals post-stroke. In: IEEE international conference on intelligent robots and systems, pp 2403–2408

148. Walters ML, Syrdal DS, Koay KL, Dautenhahn K, Te Boekhorst R (2008) Human approach distances to a mechanical-looking robot with different robot voice styles. In: Proceedings of the 17th IEEE international symposium on robot and human interactive communication, RO-MAN, pp 707–712

149. Walters ML, Koay KL, Syrdal DS, Dautenhahn K, Te Boekhorst R (2009) Preferences and perceptions of robot appearance and embodiment in human–robot interaction trials. Adaptive and emergent behaviour and complex systems—proceedings of the 23rd Convention of the society for the study of artificial intelligence and simulation of behaviour, AISB 2009, Edinburgh, pp 136–143

150. Wang E, Lignos C, Vatsal A, Scassellati B (2006) Effects of head movement on perceptions of humanold robot behavior. In: HRI 2006: Proceedings of the 2006 ACM conference on human–robot interaction, vol 2006, pp 180–185

151. Wilson EO (2006) The creation: an appeal to save life on earth. Wiley Online Library, Hoboken

152. Wittenbrink B, Judd C, Park B (2001) Spontaneous prejudice in context: variability in automatically activated attitudes. J Pers Soc Psychol 81(5):815–827

153. Yamamoto M (1993) Sozzy: a hormone-driven autonomous vacuum cleaner. In: International Society for Optical Engineering, vol 2058, pp 212–213

154. Yogeeswaran K, Dasgupta N (2014) The devil is in the details: abstract versus concrete construals of multiculturalism differentially impact intergroup relations. J Pers Soc Psychol 106(5): 772–789

155. von Zitzewitz J, Boesch PM, Wolf P, Riener R (2013) Quantifying the human likeness of a humanoid robot. Int J Soc Robot 5(2): 263–276

156. Złotowski J, Strasser E, Bartneck C (2014) Dimensions of anthropomorphism: from humanness to humanlikeness. In: Proceedings of the 2014 ACM/IEEE international conference on human–robot interaction, HRI '14. ACM, New York, pp 66–73

Jakub Złotowski is a Ph. D candidate at the HIT Lab NZ of the University of Canterbury, New Zealand and a cooperative researcher at ATR (Kyoto, Japan). He received MSc degree in Interactive Technology from the University of Tampere (Finland) in 2010. He previously worked as a research fellow at the University of Salzburg (Austria) on an EU FP7 project—Interactive Urban RObot (IURO). His research focus is on anthropomorphism and social aspects of Human-Robot Interaction.

Diane Proudfoot is Associate Professor of Philosophy at the University of Canterbury, New Zealand and Co-Director of the Turing Archive for the History of Computing, the largest web collection of digital facsimiles of original documents by Turing and other pioneers of computing. She was educated at the University of Edinburgh, the University of California at Berkeley, and the University of Cambridge. She has numerous print and online publications on Turing and Artificial Intelligence. Recent examples include: 'Rethinking Turing's Test', *Journal of Philosophy*, 2013; 'Anthropomorphism and AI: Turing's much misunderstood imitation game', Artificial Intelligence, 2011; and with Jack Copeland, 'Alan Turing, Father of the Modern Computer', *The Rutherford Journal*, 2011.

Kumar Yogeeswaran is a Lecturer of Social Psychology (equivalent to Assistant Professor) at the University of Canterbury, New Zealand. He earned his PhD in Social Psychology at the University of Massachusetts Amherst. His primary research examines the complexities of achieving national unity in the face of ethnic diversity, while identifying new strategies that help reduce intergroup conflict in pluralistic societies. His secondary research interests lie in the application of social psychology to the fields of law, politics, communication, and technology.

Christoph Bartneck is an associate professor and director of postgraduate studies at the HIT Lab NZ of the University of Canterbury. He has a background in Industrial Design and Human-Computer Interaction, and his projects and studies have been published in leading journals, newspapers, and conferences. His interests lie in the fields of Social Robotics, Design Science, and Multimedia Applications. He has worked for several international organizations including the Technology Centre of Hannover (Germany), LEGO (Denmark), Eagle River Interactive (USA), Philips Research (Netherlands), ATR (Japan), Nara Institute of Science and Technology (Japan), and The Eindhoven University of Technology (Netherlands). Christoph is a member of the New Zealand Institute for Language Brain & Behavior, the IFIP Work Group 14.2 and ACM SIGCHI.