

4

Design

What is covered in this chapter:

- How a well-designed robot can lift interactions to the next level (physical design).
- How people do not treat robots as an assembly of plastic, electronics, and code but rather as humanlike entities (anthropomorphism).
- How HRI research draws on psychological theories, such as anthropomorphism, to design and study people’s interactions with robots.
- Design methods and prototyping tools used in human–robot interaction.

How does a pile of wires, motors, sensors, and microcontrollers turn into a robot that people will want to interact with? Although it sounds like magic, the trick of turning metal and plastic into a social interaction partner is in the iterative and interdisciplinary process of robot design.

This chapter starts by exploring some general design principles and considerations (Section 4.1) before moving on to anthropomorphic design specifically in Section 4.2. In Section 4.3, different methods of designing are discussed, and Section 4.4 covers the different approaches to testing and prototyping the design you came up with. The impact of culture on HRI design is discussed in Section 4.5, and Section 4.6 wraps up this chapter by highlighting the ethical and philosophical considerations that come into play when designing a robot.

Robot design is a fast-growing field of research and practice in human–robot interaction (HRI), and the need to develop robots that are able to interact with people challenges the existing ways of designing robots. Often, robots are developed by engineers, and their interaction abilities are then tested by social scientists. This process of design starts from the inside and builds up to the outside—solving technical issues first and designing the robot’s appearance and behavior to fit. For example, a mobile platform such as a TurtleBot (see Figure 4.1) might be used as a starting point, with the desired sensors and actuators added to the body later on. If time allows, a casing could be designed to cover up all the technology. The robot’s appearance and specific social interaction capabilities then have to be built on top of this technical infrastructure. This common approach to robot building is also known as the “Frankenstein approach.” In this method, we take whatever technology

Figure 4.1 A
TurtleBot2
(2012–present)
platform. (Source:
Photo provided by
Yujin Robot)



is available and put it together to obtain a set of specific robotic functions. Clearly, such an approach is suboptimal because it commonly fails to consider a human-centered perspective that also takes into account the impact of the context and the envisioned use case.

Therefore, it is important to complement a purely technological development perspective with more holistic approaches to robot design. That is, it actually matters to consider the needs, values, and preferences of potential stakeholders and end users early on in the design process. It matters where these end users use the robot and for what purpose. Based on the characteristics of the users and the context of use, one can then decide on specific robot design features, such as appearance, interaction modalities, and level of autonomy. This might be termed a more “outside-in” mode of developing robots, in which the design process starts from the interaction that we expect the robot to be engaged in, which will determine its outside shape and behaviors. Once the design has been settled on, we work all the technology into it. Many commercial social robots are designed, at least to some degree, from the “outside in”—considering the users and how they might interact with a person and selecting or even developing technology appropriately. Honda’s ASIMO, for example, was chosen to be smaller in size so that it would not be intimidating to users. Pepper was initially designed to interact with shop visitors in Japan and has a hinged waist that allows it to bow to them as a greeting. The seal-like robot Paro was designed to inspire petlike interaction and was initially shaped like a cat, but its design was changed to a seal to address critiques users had due to their familiarity with how real cats behave; at some point in its iterative design, it also had wheels to be able to move around on the floor, but these were removed because the older adults who were its main users often had limited mobility.

Designers are trained to approach the design of artifacts in this way (see [Figure 4.2](#) for an example) and are able to make valuable contributions (Schonenberg and Bartneck, 2010). Their contribution is not limited to only the aesthetics of the robot; designers also have the skill to create



Figure 4.2
Mythical robots
designed from the
outside to the inside.
First, the shape of
the robots was
sculptured before
fitting the
technology into it.

thought-provoking robots that challenge our understanding of the roles of humans and robots.

This form of robot design often requires incorporating expertise from several disciplines—for example, designers might work on developing specific concepts for the design, social scientists may perform exploratory studies to learn about the potential users and context of use, and engineers and computer scientists need to communicate with the designers to identify how specific design ideas can be realistically instantiated in working technology (Šabanović et al., 2014). HRI design can take advantage of existing robots, designing specific behaviors or use tasks for them that fit particular applications, or it can involve the development of new robot prototypes to support the desired interactions. In either case, HRI design both takes advantage of existing design methods and develops new concepts and methods specifically suited to the development of embodied interactive artifacts (i.e., robots).

4.1 Design in HRI

4.1.1 Robot morphology and form

A common starting point for designing HRI is to think of what the robot is going to be doing. There is a debate about whether form follows function, in which the shape of an object is largely determined by its intended function or purpose, or if the reverse holds true. In HRI, clearly, form and function are inherently interconnected and thus cannot be considered separately.

Contemporary HRI designers have several different forms of robots to choose from. Androids and humanoids most closely resemble humans in appearance, but they have a lot to live up to in terms of capabilities. Zoomorphic robots are shaped like animals with which we are familiar (e.g., cats or dogs) or like animals that are familiar but that we do not typically interact with (e.g., dinosaurs or seals). HRI designers, eager to make robot appearances

Figure 4.3 Zoomorphic and minimalistic robots. From left to right: Muu (2001–2006), Keepon (2003–present), and Naked Invisible Guy. (Source: Keepon photo from Hideki Kozima, Tohoku University, ASIMO from Honda)

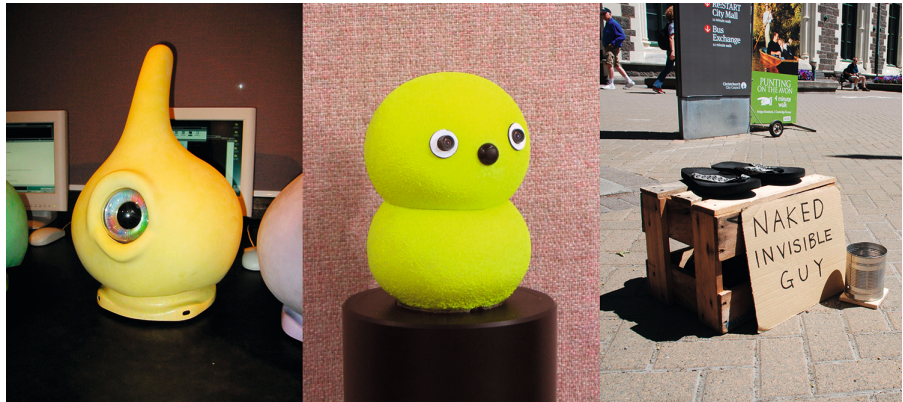


Figure 4.4 Sociable Trash Box robots are an example of objects—robotic objects with interaction capabilities. (Source: Michio Okada)



commensurate with their limited capabilities, also often design minimalist robots, which explore the minimal requirements necessary for inspiring social HRI, such as Muu (see [Figure 4.3](#), left) or Keepon (see [Figure 4.3](#), middle). The arguably most minimalistic robot is the busker robot, which consisted of a pair of animated sandals on top of a box with a signpost in front of it proclaiming “Naked Invisible Guy” (Partridge and Bartneck, 2013) (see [Figure 4.3](#), right).

Recently, the HRI field has started considering “robjects,” interactive robotic artifacts whose design is based on objects rather than living creatures, for example, a robotic ottoman (Sirkin et al., 2015), social trash cans (see [Figure 4.4](#)), or robotic toy boxes (Fink et al., 2014). Because the design space of robots is relatively large and considers questions regarding form, function, level of autonomy, interaction modalities, and how all those fit with particular users and contexts, an important aspect of design is figuring out how to make appropriate decisions about these various design aspects.

4.1.2 Affordances

The notion of affordances represents an important concept in design. This notion was initially developed as a concept in ecological psychology (Gibson, 2014), where it referred to the inherent relationship between an organism and

its environment. For example, a stone can be picked up by us and thrown away, but to a mouse, it can serve as a hiding place. The stone “affords” different interactions. This concept was amended by Don Norman 2008 to describe the perceivable relationships between an organism and its environment that enable certain actions (e.g., a chair is something to sit on, but so is a stair).

A designer needs to design a product while making its affordances explicit. Furthermore, the designer needs to incorporate user expectations and cultural perceptions. For Norman (2008), these “design affordances” are also an important way to develop common ground between robots and humans so that people can understand the robot’s capabilities and limitations and adapt their interaction accordingly. A robot’s appearance is an important affordance because people tend to assume that the robot’s capabilities will be commensurate with its appearance. If a robot looks like a human, it is expected to act like a human; if it has eyes, it should see; if it has arms, it should be able to pick up things and might be able to shake hands. Another affordance can be the robot’s interaction modalities. If a robot speaks, for example, saying, “Hello,” people will also expect it to be able to understand natural language and carry on a conversation. If it expresses emotions through facial expressions, people might expect it to be able to read their emotions. Other robotic affordances can be based on technical capabilities; for example, if it has a touch screen on its body, people might expect to interact with the robot through the touch screen. Because robots are novel interaction partners, the affordances used by designers are particularly important for signaling appropriate ways of engaging with them.

4.1.3 *Design patterns*

Because the focus of HRI is the relationship between humans and robots, the task of HRI design is not only to create a robotic platform but also to design and enable certain interactions between people and robots in various social contexts. This suggests that the main units of design that need to be considered are not only the characteristics of individual robots (e.g., appearance, sensing abilities, or actuation) but also what Peter Kahn calls “design patterns” in HRI, inspired by Christopher Alexander’s idea of design patterns in architecture (Kahn et al., 2008). Such patterns describe “a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice” (Alexander, 1977, p. x).

Within HRI, Kahn et al. (2008) suggest that patterns should be abstract enough that you can have several different instantiations, that they can be combined, that less complex patterns can be integrated into more complex patterns, and that they serve to describe interactions with the social and physical world. For example, the didactic communication pattern (where the robot assumes the role of a teacher) could be combined with a motion pattern (where the robot initiates a movement and aligns it with the human counterpart of the interaction) to create a robotic tour guide. Kahn et al.

suggest that HRI design patterns can be developed based on observation of human interactions, prior empirical knowledge about humans and robots, and designers' experiences with HRI, through an iterative design process. Some patterns they developed and have used in their designs are things like the "initial introduction" of the robot, or "in motion together," where the robot moves along with the person. Although Kahn et al.'s design patterns are not meant to be exhaustive, they emphasize the idea that the design should focus on the relationship between humans and robots.

4.1.4 Design principles in HRI

When combining the two ideas of design affordances and patterns in the process of HRI design, the usual design types that robots may be divided into, such as androids and humanoids, zoomorphic robots, minimally designed robots, or robjects, are no longer the main design focus or question. Instead, designers consider how different robot forms and capabilities fit into or express particular HRI design patterns and how they can be designed as affordances that appropriately signal the robot's interaction capabilities and purpose. With this in mind, HRI researchers have suggested some of the following principles to consider when developing the appropriate robot forms, patterns, and affordances in HRI design.

Matching the form and function of the design: If your robot is humanoid, people will expect it to do humanlike things—talk, think, and act like a human. If this is not necessary for its purpose, such as cleaning, it might be better to stick to less anthropomorphic designs. Similarly, if it has eyes, people will expect it to see; if it talks, they will expect it to be able to listen. People can also be prompted to associate specific social norms and cultural stereotypes with robots through design; for example, researchers have shown that people might expect a female robot to be more knowledgeable about dating or that a robot made in China would know more about tourist destinations in that country (Powers et al., 2005; Lee et al., 2005)

Underpromise and overdeliver: When people's expectations are raised by a robot's appearance or by introducing the robot as intelligent or companion-like, and those expectations are not met by its functionality, people are obviously disappointed and will negatively evaluate the robot. Sometimes these negative evaluations can be so serious that they affect the interaction. To avoid such problems, it is better to decrease people's expectations about robots (Paepcke and Takayama, 2010), which might have been increased by how robots are portrayed in society, as described in the "Robots in Society" chapter (see Chapter 12). This might even include not calling your design a robot because the word itself often connotes quite advanced capabilities to members of the public.

Interaction expands function: When confronted with a robot, people will, in effect, fill in the blanks left open by the design depending on their values, beliefs, needs, and so on. It can thus be useful, particularly for robots with limited capabilities, to design them in a somewhat open-ended way. This allows

people to interpret the design in different ways. Such an open-ended design approach has worked particularly well with, for instance, the seal-like robot Paro (see [Figure 2.8](#)). This baby seal robot invokes associations with pets that people have had, but it also does not get compared to animals they know, such as cats and dogs, which would inevitably lead to disappointment. As a consequence, Paro becomes a natural part of the interactions with humans and passes as a petlike character even though its capabilities are significantly below those of a typical domestic animal or that of an actual seal baby (Šabanović and Chang, 2016).

Do not mix metaphors: Design should be approached holistically—the robot’s capabilities, behaviors, affordances for interaction, and so forth should all be coordinated. If you design a humanlike robot, people may find it disturbing if it has skin covering only some parts of its body. Similarly, if the robot is an animal, it may be strange for it to talk like an adult human or try to teach you mathematics. This is related to the uncanny valley theory (see p. 66) because inappropriately matched abilities, behaviors, and appearance often lead to people having a negative impression of the robot.

Take a look at the two pictures in [Figure 4.5](#). How do they make you feel? Although both of these android representations of the science-fiction writer Philip K. Dick are perhaps a bit strange and uncanny, the one that seems unfinished and shows the robot’s insides also mixes design metaphors—the robot is both humanlike and machinelike, making it even more disturbing.

Like Kahn et al.’s (2008) design patterns, these design principles are not exhaustive but are meant to inspire thinking about how to approach



Figure 4.5 Philip K. Dick Robot (2005; rebuilt in 2010).

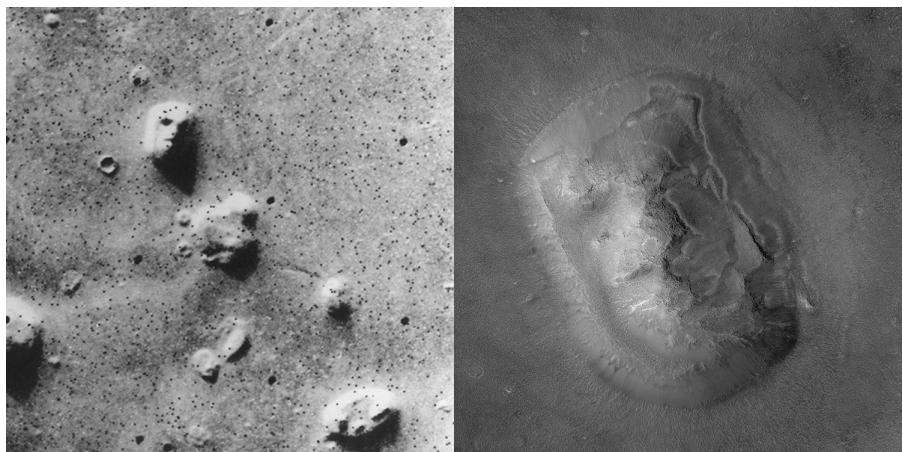
designing HRI in a way that acknowledges and incorporates the interdependence between human and robot capabilities, the need for interaction partners to be intelligible to and support each other, and the effects of the context of interaction on its success.

4.2 Anthropomorphization in HRI design

Have you ever found yourself yelling at your computer because it suddenly crashes while you are working on an essay that is due in just a few hours? You urge the computer to please bring it back again after restarting, gently touching the mouse after realizing that, indeed, the file reopens, and you can continue. You sigh in relief because “Genius”—that’s what you call your computer when no one is around to hear you—did not let you down. In fact, what you have pictured now is an ordinary scenario of a person humanizing an object, anthropomorphizing it. What a tongue twister. But what is it about, in fact?

Anthropomorphization is the attribution of human traits, emotions, or intentions to nonhuman entities. It derives from *ánthrōpos* (meaning “human”) and *morphē* (meaning “form”) and refers to the perception of human form in nonhuman objects. We all experience anthropomorphism in our daily lives. “My computer hates me!”; “Chuck [the car] is not feeling well lately”; “That grater looks like it has eyes”—you’ve either heard or uttered sentiments like this before. The latter is a special example of anthropomorphization called *pareidolia*, the effect of seeing humanlike features in random patterns or mundane objects. When the *Viking 1* spacecraft took a photo of the Cydonia area on Mars on July 25, 1976, many people saw a face on Mars’s surface, which sparked many speculations about the existence of life on Mars (see [Figure 4.6](#)). The National Aeronautics and Space Administration (NASA) sent its Mars Global Surveyor to the exact same location in 2001 to take higher-resolution photos under different lighting conditions, which revealed that the structure photographed in 1976 is certainly not a human face.

Figure 4.6 The face on Mars is an example of pareidolia. On the left is the photo from 1976, and on the right is the same structure photographed in 2001.



We will discuss anthropomorphization and anthropomorphism, respectively, in some detail as a case study of a specific design theme in HRI that incorporates technical development, psychological study, and design to enable social HRI. A robot's level of human-likeness is one of the main design decisions that robot designers need to take into account because it influences not only the robot's appearance but also the functionality it needs to offer and the social perceptions that are elicited by both form and function. In [Chapter 8](#), we will go deeper into the psychological theories underlying anthropomorphism and the consequences for impression formation.

4.2.1 *Attributing humanlike characteristics to robots*

People's innate predisposition to anthropomorphize the things around them has become a common design affordance for HRI. In anthropomorphic design, robots are constructed to have certain humanlike characteristics, such as appearance, behavior, or certain social cues, that inspire people to see them as social agents. At one extreme, android robots are designed to be as humanlike as possible; some have been fashioned as exact replicas of living humans, like a moving Madame Tussaud's wax figure (see, for example, Geminoid in [Figure 4.7](#)), or as representations of aggregated human features (e.g., Kokoro, depicted on the far right in [Figure 4.8](#)). Humanoid robots use a more abstract notion of human-likeness in their anthropomorphic designs. ASIMO (second from the right in [Figure 4.8](#)), for example, has a human body shape (two arms and legs, a torso, and a head) and proportions, but it does not have eyes. Rather, its head resembles an astronaut's helmet. Nao (see [Figure 4.8](#), middle) similarly has a humanlike body, as well as two light-emitting diode (LED) eyes that can change in color to connote different expressions, but no mouth. Some other humanoids, such as Robovie, Wakamaru (second from the left in



Figure 4.7 The Geminoid HI 4 robot (2013), a replica of Hiroshi Ishiguro. (Source: Hiroshi Ishiguro)

Figure 4.8 People readily anthropomorphize all kinds of robots, with appearances ranging from minimalist to indistinguishable from the human form. From left to right: Keepon (2003–present), Wakamaru (2005–2008), Nao (2008–present), ASIMO (2000–2018), and Kokoro’s Actroid (2003–present) android. (Source: Keepon from Hideki Kozima, Tohoku University, ASIMO from Honda)

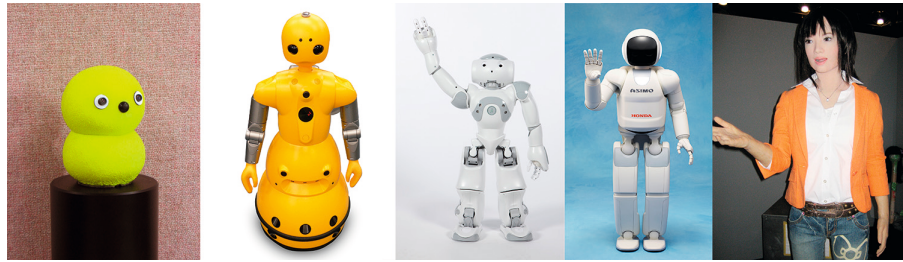


Figure 4.8), and Pepper, are not bipedal but have arms and have heads with two eyes.

Nonhumanoid robots, however, may also have anthropomorphic features. The minimalist robot Keepon (see [Figure 4.8](#), far left) has two eyes and a symmetrical body, and it likewise features displays of behavioral cues for attention and affect that may elicit anthropomorphization. Google’s autonomous car prototype has an almost cartoon-like appearance, with wide-set headlights and a button nose that suggest an anthropomorphic appearance. Festerling and Siraj (2022) also discussed the role of anthropomorphization for digital voice assistants.

Human-likeness has been key to animation designers for some time, only relatively recently sparking the interest of social psychologists. Disney’s *Illusion of Life* (Thomas et al., 1995) has inspired several social robotic projects, such as Wistort et al.’s Tofu, which displays the animation principles of “squash” and “stretch” (Wistort and Breazeal, 2009), and Takayama et al.’s work with the PR-2 using animation to give the robot apparent goals, intentions, and appropriate reactions to events (Takayama et al., 2011). Animation principles such as anticipation and exaggerated interaction have also been applied to robot design, for example, in Guy Hoffman’s Marimba player (Hoffman and Weinberg, 2010) and music companion robots (Hoffman and Vanunu, 2013). Researchers at the Honda Research Institute based the movement design of their robot Haru ([Figure 4.9](#)) on emotive actions acted out by human performers. These anthropomorphic designs take advantage not only of appearance and form but also of behavior in relation to the environment and other actors to evoke ascriptions of human-likeness.

Human-likeness in robot design includes factors related to form and appearance as well as factors relating to behavior; it may also result in the attribution of characteristics (e.g., emotions, intentions, mind perceptions) that might not be directly observable. The latter is called *psychological anthropomorphism* (Epley et al., 2007). We cover this topic in greater detail in [Chapter 8](#).

Figure 4.9 Honda Research Institute’s Haru robot.



The uncanny valley

Mori (1970) made a prediction about the relationship between the human-likeness of robots and their likability (see [Figure 4.10](#)). The idea is that the more humanlike robots become, the more likable they will be, until a point

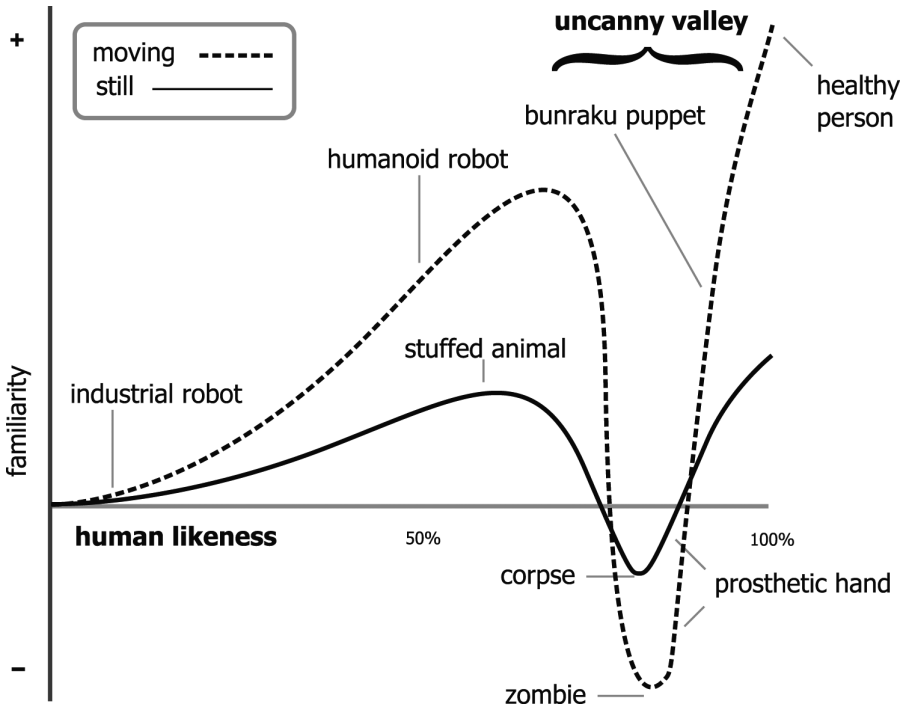


Figure 4.10 Mori's uncanny valley theory.

where they are almost indistinguishable from humans, at which point their likability decreases dramatically. This effect is then amplified by the ability of the robot to move.

Mori et al. (2012) translated Mori's original paper into English in collaboration with Mori himself. It is important to note that Mori only proposed this idea and never did any empirical work to test it. Moreover, Mori used the term 親和感 (*shinwa-kan*) to describe one of his key concepts. The translation of this concept to English remains challenging—it has been translated as likability, familiarity, and affinity. Other researchers have approached the problem by asking participants about the eeriness of the robot instead.

Unfortunately, Mori's theory has been used and abused to explain a huge number of phenomena without proper justification or empirical backup. It is often used to explain why certain robots are being perceived unfavorably, without studying the exact relationship between the features of the robot at hand and its likability. Anthropomorphism is a multidimensional concept, and reducing it to just one dimension does not model reality adequately. Moreover, the more humanlike robots become, the greater the risk of getting a certain aspect of their appearance or behavior wrong and thereby decreasing the level of likability (Moore, 2012). A simple possible explanation of why humanlike robots are liked less than, for example, toy robots is that the difficulty of designing a robot to perform to user expectations increases with its complexity.

4.2.2 Designing humanlike appearance

Robot designers may treat humanlike appearance as a characteristic of the robot itself, whereas social scientists see anthropomorphism as something that a person attributes to the robot. Considering both of these together suggests that anthropomorphism is about the relationship between robot design and functions and people's perceptions of robots.

Design approaches

To trigger anthropomorphic inferences, robot designers can take into account the dimensions of robot appearance and behavior, among many other aspects. By exploiting these aspects, they can achieve an immediate perception of the robot as more or less humanlike.

Robot appearance Graphical illustration shows us that often, only a few lines on a sheet of paper are needed to evoke the human form. In the same manner, anthropomorphism in robots can be very simple: just having two dots suggesting eyes and a simple nose or mouth is sufficient to suggest the robot is humanlike. This can be further enhanced by adding more human features, such as arms or legs, but these do not necessarily do very much to further increase the anthropomorphization. Although there are many reasons why robots look increasingly humanlike, anthropomorphization can be achieved with only a minimal set of humanlike features. Whereas androids mimic human appearance in most ways, simple robots such as Keepon and R2D2 are already very effective at triggering people to anthropomorphize. Thus, a large body of research has documented how minimal design cues might be sufficient to elicit a humanlike perception.

Robot behavior A second approach to increasing anthropomorphization is to design the behavior of an artifact such that people perceive humanlike characteristics in its behavior. Heider and Simmel (1944) showed how simple geometric shapes—triangles and circles—moving against a white background evoked people to describe their interactions in terms involving social relationships (e.g., these two are friends; this one is the attacker) and humanlike feelings and motivations (e.g., anger, fear, jealousy). Animators understand how motion, rather than form, can be extremely powerful for expressing emotions and intents. A surprisingly wide range of humanlike expressive behavior can be communicated through movement alone, without the need for humanlike form.

The Dot and the Line: A Romance in Lower Mathematics is a 10-minute animation film by Chuck Jones, based on a short book by Norton Juster. It tells the story of the amorous adventures of a dot, a line, and a squiggle. Even though the visuals are minimal, the viewer has no problem following the story. It is a prime example of how motion rather than form can be used to communicate character and intent.

Robot builders can actively encourage anthropomorphization. One effective method is to increase the reaction speed of the robot to external events: a robot that immediately responds to touch or sound will be perceived as more anthropomorphic. Such *reactive behavior*, in which the robot responds quickly to external events, is an easy approach to increase anthropomorphization. The robot jolting when the door slams shut or looking up when touched on the head immediately conveys that it is both alive and responsive. *Contingency*, responding with behavior that is appropriate for the context of the interaction, can also be used to enhance anthropomorphization. When a robot detects motion, for example, it should briefly look toward the origin of the movement. If the event, such as a tree swaying in the wind, is irrelevant to the robot, it should look away again, but if it is relevant, such as a human waving hello to engage the robot in interaction, the robot should sustain its gaze.

Although robot developers will often prefer a combination of both form and behavior to inspire users to anthropomorphize their robots, certain types of robots may be limited in how humanlike they can be. Android robots, which appear virtually identical to people, are still technically limited in their behavioral repertoire. On the other hand, developers of toy robots are often under pressure to make the hardware as cheap as possible and thus opt for an effective combination of simple visual features and reactive behaviors. It is important to also take people's expectations into account; the more apparently humanlike the robot, the more people will expect in terms of humanlike contingency, dialogue, and other features.

Impact of context, culture, and personality

People's perceptions of anthropomorphic robot design are often affected by contextual factors. Some people are more likely than others to anthropomorphize things around them, and this can affect how they perceive robots, as previous research has shown (Waytz et al., 2010). A person's demographics and cultural background can also affect their likelihood of anthropomorphizing or their interpretation of the robot's social and interactive capabilities (Wang et al., 2010; Spatola et al., 2022).

The context in which the robot is used, furthermore, can support anthropomorphization. In particular, just putting a robot in a social situation with humans seems to increase the likelihood that people will anthropomorphize it. The collaborative industrial Baxter robot, when used in factories alongside human workers, was regularly anthropomorphized by them (Sauppé and Mutlu, 2015). Furthermore, it seems that people who work alongside robots prefer them to be designed in more anthropomorphic ways: people preferred that Roomba have the ability to display its emotions and intentions with a doglike tail (Singh and Young, 2012). Workers using Baxter put hats and other accessories on it and wanted it to be more polite and chitchat with them (Sauppé and Mutlu, 2015). Workers in a car plant using a co-bot, which was named Walt and had been designed to have a blend of social features and features reminiscent of a vintage car (see [Figure 11.16](#)), considered the robot

to be a team member (El Makrini et al., 2018). Office workers who were given a break management robot gave it names and requested that it be more socially interactive (Šabanović et al., 2014).

Seeing other people anthropomorphize robots may suggest that humanizing nonhuman entities represents socially desirable behavior. To illustrate, researchers found that older adults in a nursing home were more likely to engage socially with Paro, the seal-like companion robot, when they saw others interacting with it like a pet or social companion (Chang and Šabanović, 2015). Clearly, anthropomorphic inferences may emerge instantly upon a first encounter and likewise become reshaped as a function of long-term interaction and acquaintance with a technical system. We will discuss this in more detail in [Chapter 8](#), which covers the psychology of how people perceive robots.

4.3 Design methods

Design in HRI spans a variety of methods inspired by practice from various disciplines, from engineering to human–computer interaction (HCI) and industrial design. Depending on the method, the starting point and focus of design may weigh more heavily on technical exploration and development or on exploring human needs and preferences, but the ultimate goal of design in HRI is to bring these two domains together to construct a successful HRI system.

The design process is often cyclical in nature, following this pattern:

1. Define the problem or question.
2. Build the interaction.
3. Test.
4. Analyze.
5. Repeat from step 2 until satisfied (or money and time run out).

4.3.1 Engineering design process

The engineering design method, as the name suggests, is commonly used in engineering. Starting from a problem definition and a set of requirements, numerous possible solutions are considered, and a rational decision is made on which solution best satisfies the requirements. Often, the function of an engineered solution can be modeled and then simulated. These simulations allow engineers to systematically manipulate all the design parameters and calculate the resulting properties of the machine. For well-understood machines, it is even possible to calculate the specific design parameters necessary to meet the performance requirements. If a new aircraft takes off for its maiden flight, engineers can be almost certain that it will fly. It is important to note, however, that they cannot be absolutely certain because the new aircraft will interact with an environment that is not completely predictable in all

its detail. Enough is understood, though, to be very sure of the macroscopic properties of the environment, allowing the engineers to design an aircraft that crosses the boundary from simulation to actual prototype without any hiccups. However, validating a solution in simulation is not always possible. The simulation might not be able to capture the real world in sufficient detail, or the number of design parameters can be so high that a complete simulation of all possible designs becomes computationally impossible because it would take a computer years to calculate how each solution performs. There have been some attempts at developing human–robot simulators (e.g., Lemaignan et al., 2014d), but simulating social interaction has turned out to be a very difficult problem.

Engineers working in HRI tried to design a robot to teach eight- and nine-year-old children what prime numbers are. They believed that the children’s learning would benefit from having a very personal and friendly robot, so they programmed the robot to make eye contact, use the child’s first name, and politely support the child during the quite taxing exercises. They compared the friendly robot against a robot in which the software to maintain engaging relations was switched off, expecting that robot to be the worse teacher. They were dumbfounded when the aloof robot turned out to be the better teacher by a large margin, showing how their preconceptions regarding robot design were firmly out of touch with the reality of using a robot in the classroom (Kennedy et al., 2015) (see Figure 4.11).

To make things even more difficult, some design problems can be ill-defined, or insufficient information is available about the requirements or the environment. In this case, designers may say that they are dealing with a “wicked design problem” (Buchanan, 1992), which has changing, incomplete,

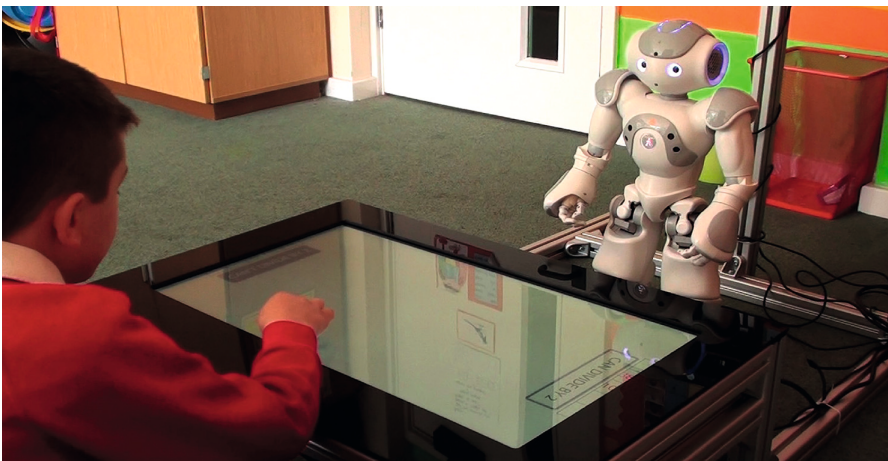


Figure 4.11 Boy learning math with a robot.

interdependent, or indeterminate requirements that make it difficult to follow a linear model of design thinking in which the problem definition can be cleanly followed by a process of problem solution. HRI design often resembles such a wicked design problem because there is a lack of information about the appropriate behaviors and consequences of robots in social contexts. Another approach to take in this case is to focus not on producing the absolute best solution but on producing satisficing solutions Simon (1996). *Satisficing* is a portmanteau of *satisfy* and *suffice*, meaning that the resulting solution will be just good enough for the purpose it is meant to serve. This is a common problem-solving approach in all human endeavors, and it is almost unavoidable in HRI, where technical capabilities may never reach the ultimate design requirement of the robot performing just as well as or better than people.

4.3.2 User-centered design process

Relying solely on the engineering design method can guide HRI development only so far, particularly when the intended uses of HRI are in open-ended interactions and spaces, outside labs or tightly controlled factory environments. In the process of satisficing, we may all too often choose not to measure the things that matter but instead only take into account what is easy to measure. One way to address this issue is to focus more specifically on the people who will use the robot and the contexts of use they inhabit throughout the design process. This can be done through user-centered design (UCD). UCD is not specific to HRI and is used in many other design domains, such as HCI, and is a broad term used to describe “design processes in which end users influence how a design takes shape” (Abrams et al., 2004). The users can be involved in many different ways, including through initial analyses of their needs and desires that can help to define the design problem, by asking them to comment on potential robot design variations to see which ones are preferable, and through evaluations of various design iterations of the robot and of the final product to evaluate its success among different users and in different use contexts.

Developers are typically confronted with having to make design decisions for which there are no obvious answers. Do people prefer the robot to have a red torso or a blue torso? Will a chirpy voice on a retail robot invite more people into the store? To answer these questions, developers often build prototypes of the different design options and test them with their target audience. By taking a human-centered perspective; considering user values, preferences, and beliefs; and running empirical evaluation studies (see [Chapter 10](#)), developers can actually ensure that the preferences or differences that they observe are not just coincidences but are really caused by the design feature under consideration. The results then inform the developers in building the best design option, and the cycle continues with new problems or design decisions. It is important to run these cycles as early as possible because the cost of making changes to the system increases dramatically later in the process. The credo is “test early; test often.”

Designers often focus mainly on the primary users—those who will mainly use the technologies. They would, for instance, investigate nurses and patients who interact with a drug-delivery robot. It is, however, also important for designers to consider secondary users. These are people who might only intermittently come into contact with the artifact or use it through an intermediary. Medical staff who see the robot in the hallway would represent an example of secondary users. Finally, the people who are affected by the use of the artifact (i.e., the tertiary users) have to be considered. These are people whose jobs might be replaced or changed as a result of the introduction of new robotic technology or who might otherwise be affected by the robot's use even if they never interact with it. These various people involved in and affected by the robot's uses are called *stakeholders*, and an initial step in the design process can involve doing some research to identify who the relevant stakeholders are. Once the stakeholders are identified, the designers can then involve them in the design process through a variety of user-centered methods, which can include needs and requirements analyses, field studies and observations, focus groups, interviews and surveys, and user testing and evaluations of prototypes or final products (Vredenburg et al., 2002). We will discuss several of these methods in [Chapter 10](#).

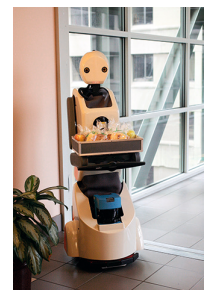
Carnegie Mellon University's Snackbot was designed through a user-centered process that involved taking into consideration the robot, people, and the context. The design process was iteratively performed over 24 months and involved research on where people could already get snacks in the building to establish need, initial technology feasibility and interaction studies, multiple prototypes, and further studies of how the robot was used and the effects of different forms of dialogue and robot behaviors on user satisfaction (Lee et al., 2009) (see [Figure 4.12](#)).

Figure 4.12
Snackbot (2010), a system developed at Carnegie Mellon University to study robots in real-world settings. (Source: Photo provided by Jodi Forlizzi)

4.3.3 Participatory design

HRI researchers increasingly use more collaborative and participatory design approaches. Both collaborative and participatory methods seek to include the potential users and other stakeholders, or people who might be affected by robots, in the process of making decisions about appropriate robot design from early on in the design process. This is clearly distinct from the notion of bringing users in at the evaluation stage, where the design is partially or fully formed and users' input is largely used to test particular factors and assumptions already expressed in the design. In this way, participatory design recognizes the expertise people have about their everyday experiences and circumstances.

Participatory design has been present in the design of other computing technologies, particularly information systems, since the 1970s, when it was used to enable workers in organizations to participate in the design of software and other technologies that they would use in their work later on. Participatory



design in HRI has been working on developing ways for users to become engaged in the process of making design decisions about robots—for instance, by testing and developing particular behaviors for robots, designing robot applications for their local environments, and conceptualizing how existing robotic capabilities can potentially address their needs and fit into their everyday contexts. DiSalvo et al. (2008) performed one of the early participatory design projects in HRI in their “neighborhood networks” project. Here, community members used a robotic prototype provided by the researchers to develop environmental sensors for their neighborhood. In another participatory project, roboticists and visually impaired community members and designers worked together in a series of workshops to develop appropriate guidance behaviors for a mobile PR-2 robot (Feng et al., 2015). Participatory design has also been used in various healthcare and educational applications for HRI (see, e.g., Šabanović et al., 2015). Teenagers (Björling et al., 2019) and even children (Zaga, 2021) have also participated in the design of HRI through various participatory design methods.

Participatory design is always challenging, but working on participatory design with robots has its particular difficulties. One is the fact that people have many different preconceptions about robots but little knowledge about the technology involved in making them, which leads to unrealistic design ideas. At the same time, designers have little knowledge of the day-to-day lives and experiences of people in many of the applications in which HRI is most needed (e.g., eldercare). While working with older adults and nursing home staff to develop assistive robots for older adults with depression, Lee et al. (2017) and Winkle et al. (2018) focused on supporting a process of mutual learning between HRI researchers and participants, which allowed both sides to explore and teach each other about their different areas of expertise. This also helped support participants’ learning to start thinking about design beyond just designing for themselves. HRI researchers have also developed frameworks to support the interdisciplinary and participatory design of social robots (Axelsson et al., 2021). Participatory design is still new in HRI, but with more and more applications being envisioned for diverse populations and everyday contexts, it is becoming an increasingly important component of the HRI design methods toolkit.

4.4 Prototyping tools

Although it is possible to develop simple robot prototypes from materials that are generally available, such as cardboard or found objects, several prototyping kits and tools for creative interactive technologies have recently become available on the market. These make it possible for a wide variety of people with different levels of technical expertise and economic resources to try their hand at robot design. They also enable more rapid and iterative development of robot designs by making the representation of interaction a simpler thing to create.



Figure 4.13 LEGO Mindstorms (1998–2022) was the brainchild of Seymour Papert, a Massachusetts Institute of Technology professor who was an avid proponent of using computers to support child learning.

Perhaps the earliest type of kit that could be used for developing different robot designs was the first-generation LEGO Mindstorms system (see [Figure 4.13](#)), which provided bricks for building and specialized bricks for programming and actuating simple robot prototypes. Bartneck and Hu (2004) used LEGO robots to illustrate the utility of rapid prototyping for HRI, and the first case studies had already appeared in 2002 (Klassner, 2002).

The Vex Robotics Design System¹ is also widely known and used, and its advanced version is the kit of choice for the popular FIRST Robotics Competitions.² More recent additions to the array of kits available are Little Bits, which provide easy-to-use plug-and-play electronic bricks, including sensors and actuators, among others, that can be used to quickly and easily create interactive prototypes.

The Arduino microcontroller³ is very affordable and has a large hobbyist community providing open-source designs and code, as well as a wide array of peripherals (sensors, motors, LEDs, wireless units, etc.) that allow for more flexibility in design but require more technical know-how.

Other equipment, such as the Raspberry Pi⁴ single-board computer and affordable and even portable three-dimensional (3D) printers, can not only make HRI prototyping easier but also may even be said to be making it accessible to the masses (or at least to college students).

Designers also incorporate other existing technologies into robot design, including smartphones. Even an average smartphone these days has sufficient computing power to control a robot. Furthermore, a smartphone has many

¹ See www.vexrobotics.com

² See www.firstinspires.org/

³ See www.arduino.cc

⁴ See www.raspberrypi.org

Figure 4.14
Robovie MR2
(2010) is a
humanoid robot
controlled through a
cell phone.



built-in sensors (microphone, camera, gyro sensor, accelerometer) and actuators (screen, speaker, vibration motor). The Robovie MR2 is an early example of integrating a smartphone into a robot to control all of its functions (see [Figure 4.14](#)). Hoffman calls this the “dumb robot, smartphone” approach to social robot design (Hoffman, 2012).

Available technologies for prototyping continue to develop, fueled at least in part by ongoing efforts to engage more students, hobbyists, and even potential users in technology design.

4.5 Culture in HRI design

As not only an interdisciplinary but also an international field of research, HRI design has been particularly interested in the question of cultural effects on perceptions of and interactions with robots. Culture, the different beliefs, values, practices, language, and traditions of a group of people, plays into robot design both in the form of factors introduced by designers and in the context in which users interpret different HRI designs.

Researchers commonly make connections between cultural traditions and the design and use of robots, particularly contrasting the norms, values, and beliefs in the East and West: animist beliefs have been used to explain the perceived comfort of Japanese and Korean populations with robots (Geraci, 2006; Kaplan, 2004; Kitano, 2006), whereas human exceptionalism has been suggested as a source of Westerners’ discomfort with social and humanoid robots (Geraci, 2006; Brooks, 2003). Holistic and dualistic notions of mind and body (Kaplan, 2004; Shaw-Garlock, 2009) and individualist and communitarian social practices (Šabanović, 2010) have been identified as design patterns represented in the design of robots and potential human interactions with them.

In addition to these generalized connections between culture and robotics, HRI researchers have been studying cultural differences in and effects on people’s perceptions of and face-to-face encounters with robots. In a comparison using Dutch, Chinese, German, U.S., Japanese, and Mexican participants, it was found that U.S. participants were the least negative toward robots, whereas the Mexican participants were the most negative. Against expectations, the Japanese participants did not have a particularly positive attitude toward robots (Bartneck et al., 2005). MacDorman et al. (2009) showed that U.S. and Japanese participants have similar attitudes toward robots, suggesting that such factors as history and religion (see [Figure 4.15](#)) may affect their willingness to adopt robotic technologies. Survey evaluations of the seal-like robot Paro by participants from Japan, the United Kingdom, Sweden, Italy, South Korea, Brunei, and the United States found that participants generally evaluated the robot positively but identified different traits as most likable according to their country of origin (Shibata et al., 2009).

In the context of human–robot teamwork, Evers et al. (2008) found that users from China and the United States responded differently to robots and that

Figure 4.15 The
BlessU2 robot was
used by the
Protestant church in
Germany to give
blessings.



human team members found robots more persuasive when they used culturally appropriate forms of communication (Lindblom and Ziemke, 2003). Findings from two generative design studies with participants in the United States and South Korea, which asked users to think about robotic technology in their own homes, showed that user expectations of and needs for robotic technologies are related to culturally variable conceptions of the home as relation oriented in Korea and more functionally defined in the United States (Lee et al., 2012). The growing body of work on cross-cultural differences in HRI and their potential design implications identifies that cultural considerations should be taken into account when designing robots, both for international and local uses.

4.6 From machines to people—and the in between

As the previous discussion shows, designing human–robot interactions involves making many decisions about the form, function, and desired effects of robots. HRI designers, however, also bring deeper philosophical, ethical, and even political commitments into their work. Although these can be unconsciously brought into HRI research, we think it is useful for HRI scholars to consciously engage with these concerns in the course of their robotics research and development.

One of the most basic decisions that robotics researchers make is the type of robot they want to work on—is it meant to resemble a human or be more like a machine? Another decision can involve the main goals of the work—is it focused on producing technical developments, understanding humans, or perhaps developing HRI systems that can be used for specific applications and contexts of use? These decisions have significance beyond just the design and use of the robot, however. One could argue that the creation of robots by their designers, in particular those in which robotic copies of actual people are created, is an immortality project. Such projects are “symbolic belief systems that promise that the individual will not be obliterated by the demise of his or her physical body” (Kaptelinin, 2018, p.6). Hiroshi Ishiguro’s work on android copies of living human persons is a case in point, in which the robotic copy can aim to stand in the place of that specific person, both in current and ostensibly future interactions. Ishiguro himself describes how he feels his own identity is interconnected with the robot, which persists as a replica of his past and younger self that he now feels the pressure to emulate (Mar, 2017). But the relationship between machinelike robots and designers can be just as deep. Describing his work with industrial robots, Japanese roboticist Masahiro Mori defined the relationship between humans and machines as being “fused together in an interlocking entity” (Mori, 1982). This close relationship has direct consequences for the form and function of the robot on the one side and the designer on the other side, as well as on the future consequences and uses of the robot in society.

Robot design can also be guided by a personal commitment to specific social and philosophical values, such as improving access to resources for

broader populations, increasing participation in the design of and decision-making about robots, or contributing to the solution of pressing social issues. Roboticist Illah Nourbakhsh described how his personal values affect his robotic projects as follows:

One way out is to say my work is purely theoretical, who cares how somebody applies it? I didn't want to do that. I wanted to say my work involves theoretical components, but I'm taking it all the way to seeing a real result in the physical world. And furthermore, I want it to be socially positive in some measure. ... I want to work on something so socially positive that not only do I hope everyone uses it, but I want to see at least one used case to fruition. Then you have this feedback loop from real-world application back to engineering design. (Šabanović, 2007, p. 79)

In this way, the choice of what type of HRI project to pursue and the goals to focus on in design can reflect personal or collective values (e.g., of the research group or of project collaborators).

Relatedly, it is not only researchers' values that matter, but likewise, a human-centered approach should take into account user and organization values, for example, in the framework of value-sensitive design (VSD) (Friedman et al., 2002). Indeed, although VSD represents an established method to advance novel technologies, it has rarely been used in the context of social robots. As a research method for social HRI, VSD can help integrate user perspectives in a literally valuable way (see also Schmiedel et al., 2022).

These authors point out that within the VSD framework, technologies adapt to human needs rather than vice versa. By means of VSD, human values can be translated into technological requirements, thereby ensuring that user or stakeholder perspectives are integrated at the onset of technology development by means of value identification, value embedding, and value evaluation.

Figure 4.16 Robert M. Pirsig (September 6, 1928–April 24, 2017) is the author of *The Metaphysics of Quality*, which has inspired many designers.



One of the authors finds inspiration for his design in the work of Robert M. Pirsig (see Figure 4.16), who put it this way:

The real [aesthetics] lies in the relationship between the people who produce the technology and the things they produce, which results in a similar relationship between the people who use the technology and the things they use. (Pirsig, 1974, p. 299)

Pirsig emphasizes the crucial role of obtaining peace of mind in order to arrive at good design as the barrier between the designer and the object to be designed dissolves:

So the thing to do when working on a motorcycle, as in any other task, is to cultivate the peace of mind which does not separate one's self from one's surroundings. When that is done successfully then everything else follows naturally. Peace of mind produces right values, right values produce right thoughts. Right thoughts produce

right actions and right actions produce work which will be a material reflection for others to see of the serenity at the centre of it all. (p. 305)

The connection between the robot and its designer is far deeper than you may assume. Pirsig spent his whole life working out *The Metaphysics of Quality*, in which he argues that there is no fundamental difference between the designer and the object he or she designs. What connects them is “quality.”

Considering the peace of mind of the designer might sound strange at first, but Pirsig argued that in the moment of the perception of quality, there is no division of objects and subjects. In the moment of such pure quality, the subject and the object are one (Pirsig, 1974, p. 299). Artists might be familiar with the experience of unity with their work, and the work of designers and engineers might be enhanced if they, too, would be more sensitive to this connection.

4.7 Conclusion

Designing robots requires multidisciplinary expertise, often by means of a team, and a process that takes the users and the interaction context into consideration. Various prototyping tools are available to quickly build and test robots. Once the users and their interactions with the robot are understood, the robot needs to be designed from the outside in—starting with the potential users and use context to develop design concepts and the technical specifications for the robot. HRI designs also express, whether consciously or unconsciously, the social and ethical values of the designers.

The robots’ anthropomorphism is one of the most important design considerations in contemporary HRI. We provided a detailed description of the construct of psychological anthropomorphism as a prime opportunity for a fruitful exchange between disciplines, leading to a broader overall understanding of the concept in the social sciences and robotics. Beyond the theoretical and methodological gains from investigating anthropomorphism, HRI studies have also shown the importance of considering humanlike form and function in robot design for perceived interaction quality, HRI acceptance, and enjoyment of the interaction with humanlike robots.

Questions for you to think about:

- Think about the features of a humanlike robot in terms of “design affordances.” Which affordances should be considered in humanlike robots?
- Try to think about “design patterns” for social robots that greet people daily. Find and describe repeatedly reused patterns in behavior.
- Imagine you have to design a robot. Consider the necessary steps, taking a participatory design approach.

- Discuss the role of user expectations in robot design. What are important points to consider if you want to market your robot?
- What is your opinion: Should a social robot have very few humanlike cues, or should it be highly anthropomorphic in design (e.g., like an android)? Which robot would be accepted more by people in general? Why?
- Think about a robot that you might want to have in the near future. Picturing this robot, try to think about a way to encourage more anthropomorphization based on its behavior. Which behaviors should the robot show to be perceived as humanlike?

4.8 Exercises

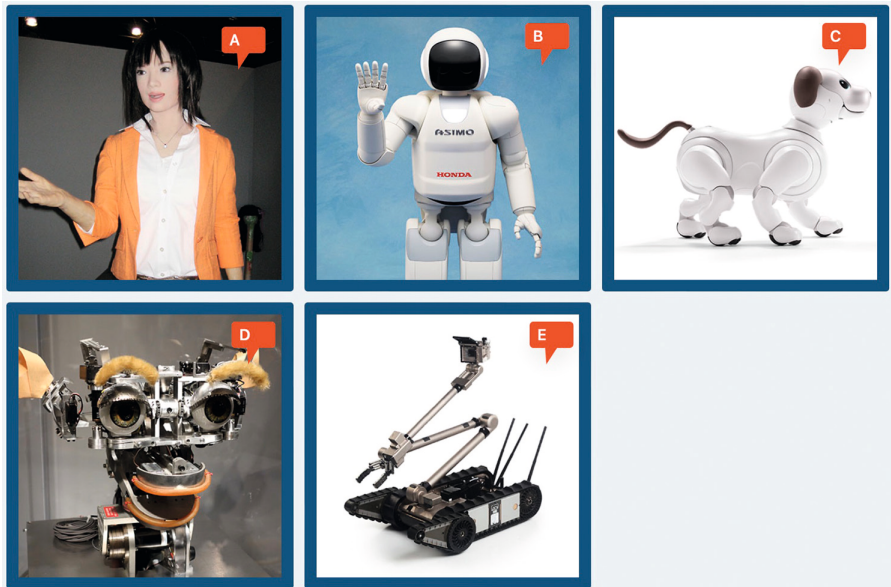
The answers to these questions are available in the Appendix.

* **Exercise 4.1 Pareidolia** Take photographs of pareidolia in your environment. Do not just google images. Use your phone or camera. Why did you choose these images?

* **Exercise 4.2 Anthropomorphism** Have a look at [Figure 4.17](#). Sort the robots from low to high anthropomorphism.

1. Lowest anthropomorphism: _____
2. Low anthropomorphism: _____
3. Medium anthropomorphism: _____
4. High anthropomorphism: _____
5. Highest anthropomorphism: _____

Figure 4.17 Different robots. (Source: B, Honda; C, Copyright of Sony Corporation)



*** **Exercise 4.3 Design an autonomous vehicle** Watch this video, and then answer the question that follows.

- Dr. Leila Takayama, “What Is It Like to Be a Robot?” <https://youtu.be/bFRBpVhqrXo>
- 1. If you were designing an automated self-driving car, like the ones developed by Google or Tesla, what kinds of affordances and/or design patterns would you include in the design to make people be and feel safe in the car as passengers and allow pedestrians and other drivers on the road to be able to trust the car in traffic? You can refer to [Chapter 1](#) and this chapter, as well as Leila Takayama’s talk (linked in the previous exercise), which discusses the sense of control in autonomous systems and some car examples, among other things, to justify your design decisions.

Future reading:

- Duffy, Brian R. Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3):177–190, 2003. ISSN 0921-8890. doi: 10.1016/S0921-8890(02)00374-3. URL [https://doi.org/10.1016/S0921-8890\(02\)00374-3](https://doi.org/10.1016/S0921-8890(02)00374-3)
- Fink, Julia. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In Ge, Shuzhi Sam, Khatib, Oussama, Cabibihan, John-John, Simmons, Reid, and Williams, Mary-Anne, editors, *Social Robotics*, pages 199–208. Springer, Berlin, 2012. ISBN 978-3-642-34103-8. URL https://doi.org/10.1007/978-3-642-34103-8_20
- Kahn, Peter H., Freier, Nathan G., Kanda, Takayuki, Ishiguro, Hiroshi, Ruckert, Jolina H., Severson, Rachel L., and Kane, Shaun K. Design patterns for sociality in human-robot interaction. In *The 3rd ACM/IEEE International Conference on Human-Robot Interaction*, pages 97–104. Association for Computing Machinery, New York, 2008. ISBN 978-1-60558-017-3. doi: 10.1145/1349822.1349836. URL <https://doi.org/10.1145/1349822.1349836>
- Lowdermilk, Travis. *User-Centered Design: A Developer’s Guide to Building User-Friendly Applications*. O’Reilly, Sebastopol, CA, 2013. ISBN 978-1449359805. URL <http://worldcat.org/oclc/940703603>
- Norman, Don. *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books, New York, 2013. ISBN 9780465072996. URL <http://worldcat.org/oclc/862103168>
- Pirsig, Robert M. *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*. Morrow, New York, 1974. ISBN 0688002307. URL <http://worldcat.org/oclc/41356566>
- Simon, Herbert Alexander. *The Sciences of the Artificial*. MIT Press, Cambridge, MA, 3rd edition, 1996. ISBN 0262691914. URL <http://worldcat.org/oclc/552080160>