PRoSPEro

Pedagogy of Robotics in the Social Professions in Europe

Scoping Paper 2
Robot Technology & Design

Dennis Reidsma, Daniel Davison, Edwin Dertien, Bob R. Schadenberg, Cristina Zaga, Lonneke Kolkman
University of Twente, the Netherlands

This scoping paper is intellectual output O2 of the PRoSPEro project
11 December 2019

Funded by the
Erasmus+ Programme
of the European Union

Pedagogy of robotics in the social professions in Europe. Project funded by the European Union under the Erasmus+ Programme Key Action 2, 2018-2019. Project 2018-1-IE02-KA203-000611

Document details

Project no	2018-1-IE02-KA203-000611
Project title	Pedagogy of robotics in the social professions in Europe
Deliverable	Scoping paper 2
Delivery date	11 December 2019
Workpackage	1
Intellectual output	O2
Author(s)	Dennis Reidsma, Daniel Davison, Edwin Dertien, Bob R. Schadenberg, Cristina Zaga, Lonneke Kolkman
Institution(s)	University of Twente, the Netherlands
Reviewer(s)	
Approved by	Coordinator
Dissemination type	Internal to project (initially) Public (website) – end project
Document type	Scoping paper
No of pages	25
No of words	9470 including bibliography

Abstract

The report is written in the context of the EU Erasmus+ PRoSPERo project, which concerns the use of social robots in the so-called "social professions" (nursing, early childhood education, child care, elderly care, etc). Social robots are increasingly entering these professions; PRoSPERo addresses the implications of this development for the education of future professionals in those domains. What do they need to learn about social robots (including their technology, design, ethics, and relevant policy and regulation aspects) in order to (a) be able to deal with those robots as they enter their professional environment (so, use them productively and experience them as a helpful addition to the workplace) and (b) be able to contribute in a meaningful way to the design of social robot applications in the domain, including choosing the direction in which such applications should be sought? Other deliverables in the project address different relevant topics for curriculum content, such as the ethics of social robots in the social professions, policy surrounding the domain, pedagogy and technology, and an overview of trends of social robots entering the social professions (including an overview of illustrative projects from related work); this report focuses on the technology and design perspective.

1 Introduction

This scoping review concerns *technology for robots that act in the social domain*.

1.1 Project context

The report is written in the context of the EU Erasmus+ PRoSPERo project, which concerns the use of social robots in the so-called “social professions” (e.g., nursing, early childhood education, elderly care). Social robots are increasingly entering these professions; PRoSPERo addresses the implications of this development for the education of future professionals in those domains [46]. What do they need to learn about social robots (including their technology, design, ethics, and relevant policy and regulation aspects) in order to (a) be able to deal with those robots as they enter their professional environment (so, use them productively and experience them as a helpful addition to the workplace) and (b) be able to contribute in a meaningful way to the design of social robot applications in the domain, including choosing the direction in which such applications should be sought?

Other deliverables in the project address other topics, such as the ethics of social robots in the social professions, policy surrounding the domain, pedagogy and technology, and an overview of trends of social robots entering the social professions (including an overview of illustrative projects from related work); this report focuses on the technology and design perspective.

1.2 Goal of this report

This report is a first step towards structuring state of the art knowledge on social robots from a technology and design perspective, in a way that contributes towards the above mentioned educational goal. The intended reader of this report are therefore people working / researching in these social professions (so: people who do not necessarily have a technology or design background). Given the current conversations on social robots that we find in media, news, and popular science, in which robots and AI are often ascribed almost mythical capabilities of intelligence, autonomy, and social skills, an important goal of this report is also *demystifying* the domain of social robots, in order to help care professionals without technology or design background take co-ownership of social robot development.

1.3 Topics covered

The report addresses a number of topics:

- definitions of what we consider social robots
- basic technologies for the physical capabilities of social robots
- basic technologies for the social capabilities of social robots
- design and use of social robots
- conclusions

2 What is a social robot?

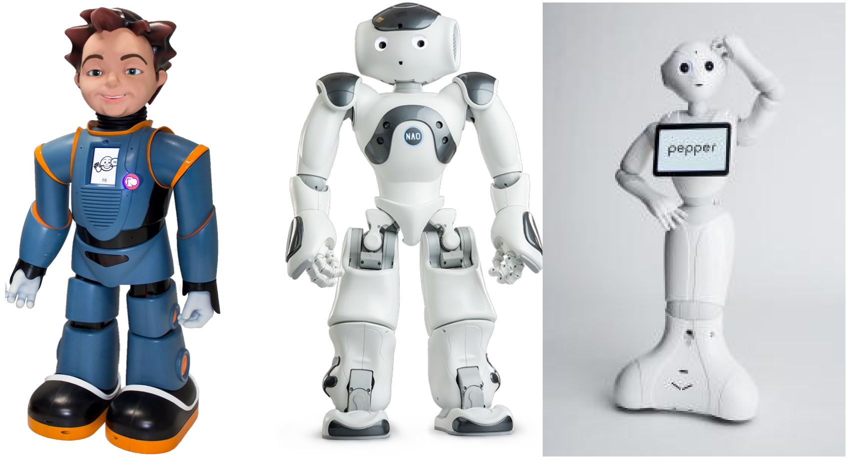
The imagery of robots that we encounter in media and popular science, and therefore the image that most people are most familiar with, is largely “puppet” oriented (robots are pets or little men with a face and eyes). However, the field of interactive social robots may be seen to encompass a much larger scope. In this report we will not draw very strict boundaries of what is or is not a social robot. A wider range of possible definitions will help social professionals to better understand the possibilities and implications of social robots; and in fact some of the discussions in this report apply as well to non-robot social and intelligent technology. As such we will work with a definition of social robots of somewhat wider scope than used by Breazeal et al. [6].

2.1 To start with: Six examples of robot platforms

To start with, this section provides a a sketch of six social robot platforms to informally span the design space. These robots will be used to exemplify the various aspects of social robots that we discuss throughout this report. Note: these are *platforms*, and not yet meant as illustrations of *applications that use the platforms*.

2.1.1 Little-man-with-face

Alice[1] (see Figure 1) is a full humanoid robot: it has arms, legs, a face; can walk, gesture, and gaze at something; and has a voice including lip movement. Such robots (Nao and Pepper[2] are other well known examples) are meant to be puppeteered with human-like expressive conversational behaviour; they are often employed in applications in which the core function either depends on social conversation, or on human like embodied expression (e.g. the robot showing how to do certain motions). The various types of humanoid robot platform differ in aspects such as the presence of a realistic face, the quality and range of body motion, the capability to grasp things with the hands, and whether and how locomotion is possible.



(a) the Alice robot

(b) the Nao and Pepper robots

Figure 1: Examples of humanoid robots.

---

[1] See https://www.doxy.nl/film/ik-ben-alice/; Also known as Zeno or Milo, see http://www.robokind.com

[2] http://softbankrobotics.com

2.1.2 Pet

Paro[3] (see Figure 2) is a small baby seal that can be held and cuddled, and responds in an emotionally expressive way to relatively simple behaviours by the human user. Other examples of Pet-robots are Pleo[4], Furby[5], the Hasbro interactive cat[6], and the MiRo robot[7]; the more restrictive expressiveness of these robots often focuses on displaying directed attention in combination with various emotions. Because these robots are more pet-like people have lower expectations of their intelligence; often they are employed to engage in emotionally expressive interaction rather than the conversational interaction of humanoid robots. More often than with humanoid robots, some of these pet robots (although not all of them) are closed platforms. For those, among which Paro and Pleo, it is not possible for third parties to develop new behaviours or interactive programs for them; one simply has to take the robot's existing behaviour and fit it into an envisioned application.

(a) the Paro baby seal

(b) the Pleo and Furby robots

Figure 2: Examples of pet robots.

2.1.3 Skype-on-wheels

The Teresa robot[8] (Figure 3; other examples are the Double robot[9] and the Giraff[10]) is a video conferencing unit on wheels. The remote person who is visible on the screen can make the robot drive around in the room, which gives him or her more *presence* in the room and the opportunity to proactively watch things from a different point of view through the robots webcam. Other people in the room do not so much interact with the robot itself, but more with the person who is visible in the videoconferencing screen of the robot. However, as discussed later, such robots are still capable of taking semi autonomous action. The robot in the Teresa project

---

[3] http://www.parorobots.com/

[4] https://www.pleoworld.com/pleo_rb/eng/index.php

[5] https://furby.hasbro.com/en-us

[6] https://joyforall.com/products/companion-cats

[7] http://consequentialrobotics.com/

[8] https://whirl.cs.ox.ac.uk/teresa/

[9] https://www.doublerobotics.com/

[10] http://www.giraff.org/

was used to investigate how the task of “socially appropriate” navigation in a room can be a joint task between the controlling user and the autonomous capabilities of the robot. Skype-on-wheels are typically built for indoor settings; some of the platforms furthermore have been built to support multi person interaction (e.g. multiple microphones plus software to filter out and transmit the sound of only the current speaker instead of all room ambient sounds around the robot)

(a) the Teresa robot (b) the Double and Giraff robots

Figure 3: Examples of Skype-on-wheels robots.

2.1.4 Embodied voices

Amazon’s series of Alexa devices[11] (See Figure 4) are a prime example of an “embodied voice”: their functionality is focused on spoken conversational intelligence. Their embodiment (the smart speaker) provides the user with a clear focus point in the conversation; it is made more prominent by lights (e.g., the Homey[12]) and sometimes the suggestion of a face or googly eyes (e.g., the Muu Socia 3.0[13]). Some “embodied voices” are purely command based, recognizing a limited set of predefined command words; other systems aim for at least the suggestion of full fledged conversational capabilities. Looks vary between more “simple speaker like” and more robot like. Finally, some of these systems specifically also aim for integration with the smart home, and allow one to access all domotics devices in the house through interaction with the embodied voice system.

(a) Alexa speakers (b) Homey and MuuSocia 3.0

Figure 4: Examples of embodied voice robots.

---

[11] https://developer.amazon.com/alexa/echo
[12] https://www.athom.com
[13] https://moosocial.com/

(a) the Relay robot (b) the Roomba robot Figure 5: Examples of Figure 5:

Figure 5: fetch-and-carry robots.

2.1.5 Fetch-and-carry

The Relay[14] (see Figure 5) is a medical delivery robot that is loaded with medical supplies which it then autonomously brings to the right department and room in a hospital. It exemplifies a class of robots that are designed to execute practical functions in the real world (in this case, delivery of packets) in a (semi) autonomous way. They happen to execute their function in a social space with people but are not primarily aimed for social interaction with users. Other typical examples are the Xenex Lightstrike Robot[15] (sterilisation of equipment and rooms on request); the Roomba (vacuum cleaner)[16], SAM the security robot [17], and LEA the robot walker support for elderly people[18]. The details of the platforms are widely different, because the functionality of the platform is so closely tied to the practical task that the robot has to carry out.

2.1.6 Wearable body augmentation

Finally, Goosebumps[19] (see Figure 6) is a piece of robotically actuated clothing that responds with movement and color changes to input from its sensors. Such wearable body augmentation[20] often uses this to express something in a visual and tactile way – for example, express to the wearer through a gentle squeezing that they should relax more (Textile Reflexes[21]) or express to others something about the mood of the wearer (Mood Sweater[22]). These systems depend on integrated sensors that perceive the (social) state and context of the user, using biophysiological and motion sensors. Many of these systems fall under artistic or speculative design, or concept prototypes; only very few have been explored as to their effect on actual people's lives.

---

[14] https://www.savioke.com/

[15] https://www.xenex.com/

[16] https://www.irobot.com/roomba

[17] https://www.sparkdesign.nl/en/projects/security-robot-sam-secure-autonomous-and-mobile

[18] https://www.robotcaresystems.com/

[19] http://sensoree.com/artifacts/awe-goosebumps/

[20] Note that these are different than e.g. exoskeletons since their primary purpose is to augment and operate in social (inter)actions.

[21] https://www.wearableroboticslab.nl/Projects/Textile Reflexes/

[22] http://sensoree.com/artifacts/ger-mood-sweater/

(a) Goosebumps (b) Textile Reflexes and Mood Sweater Figure 6: Examples of Figure 6:

Figure 6: wearable body augmentations.

2.2 A first sweeping definition: how are these platforms the same?

The previous section presented examples of social robots in a broad interpretation of the word. We base our first sweeping definition of how a system could be classified as social robot on three pragmatic viewpoints.

A social robot operates in a "social space". A social space contains people who interact with the system and/or each other and who generally behave socially; the social nature of the space and the social nature of the activities in it are relevant to the functioning of a social robot.

A social robot has a physical embodiment, even if it is only a smart speaker with googly eyes. Although many of the things we discuss are also relevant for virtual agents, we nevertheless focus on social robots that act in the real world. Note that the "body" of a robot is not just a container for its sensors and actuators. It is also a physically present embodiment of its social agency: literally the focus of interaction between the user and the socially intelligent system. In various contexts, this physical presence has been investigated for its associated positive effects [31, 28, 32].

A social robot operates on the sense–think–act paradigm. They perceive and interpret the physical and social space around them, follow rules and patterns to decide how to respond, and have a repertoire of practical and/or expressive behaviours available to carry out this response. The sensing, thinking, and acting could be said to form three aspects of the social "intelligence" of the robot; any specific social robot may have extensive intelligence in all three, two, or even only one of these domains (e.g., the perception of a robot vacuum cleaner is quite limited compared to that of a complex conversational robot).

2.3 Adding nuance: how are these platforms different?

It is clear from the descriptions of the six example robot platforms that they range across a spectrum of how exactly they are social, and how exactly they can be seen to be a robot. Below, we set out some of the most visible dimensions on which the platforms can differ from each other. The combination of this and the previous section will also set out the topics that we will further discuss in the rest of the document.

Form The first and most visible factor that distinguishes the six platforms in Section 2.1 is their type of embodiment, ranging from humanoid characters to robotically actuated clothing. For the fetch-and-carry and skype-on-wheels types of robot, form mostly follows function. For the other types, the form *is* the function, defining its associated social expressiveness and its cuteness/furryness. A major factor in choosing the embodiment relates to the *realism* of form and motion, and to the extent of underpromising versus overpromising of the social capabilities of the robot. In that sense, the spectrum of humanoid and the simpler pet like robots can be extended with what we call “minimal expressive robots”: robots that only offer the minimal necessary affordances needed to express something: e.g. an asymmetric form that only rotates in place to indicate direction of attention and awareness of humans. Cristina Zaga coins the term “Robothings” [56] for robots that have a thing-like-appearance, role and behavior. Zaga et al. designed the robothing Push-one, an omnidirectional remotely controlled robot that plays with children in puzzle games and only interacts by locomotion and simple movements: pushing pieces around and rotating on its base (see Figure 7). They used this, among other things, to explore the potential for minimally expressive robots to communicate meaningfully and expressively through (physically) limited capabilities [55]. This also fits the view of Taipale et al (2015), who seek to move beyond what they call ’human-mimicking’ machines, suggesting that the social context of interaction between machine and human is sometimes more significant than the physical form and actions of the robot.

Figure 7: The Push-one minimally expressive robot.

Technological complexity The next factor on which the robots quite obviously vary is their technological complexity. This concerns the complexity of their sensors, actuators and navigation, but also the AI, social sensitivity, and dialogue capabilities. On the relatively low end of technological complexity is the typical “plush pet robot” (the Hasbro interactive cat; Furby). Embodied voices are typically low on sensor, actuator and navigation technology, but score often high on AI and conversational technology; fetch-and-carry robots often score the other way around (complex sensors, actuators and navigation, but low conversational capabilities). Little-man-with-a-face robots, finally, score often high on both types of technological complexity. Note, however that robots at the low end of technological complexity can still be highly social, as long as they are responsive and expressive in exactly the right way. This is exemplified by Zaga’s minimally expressive robot Push-one, but also, for example, by Paro that has only a few expressive capabilities but shows a rich and highly targeted responsiveness in its limited functions, leading to high social capabilities. We will go deeper into these topics in the later sections on dialogues, the basic technologies of autonomy, and the section on AI.

Socialness as prime purpose Thirdly, social robots differ in the extent to which their socialness is part of their prime purpose, in the core of its functionality. For example, for comfort & companionship robots having social interaction is their main purpose. For other robots, their primary function may be practical (vacuum cleaning), and their need for acting socially appropriate stems indirectly from the fact that they operate in a social context that may impact their functioning. One can still actively take this into account in their design: for example, a medical delivery robot, when overtaking multiple people in a corridor, may take into account whether these people are a family group or unconnected strangers when deciding how exactly to overtake them.

Socialness as designer's intent or user's attribution A variation on the previous dimension is the question of whether the socialness of the robot is a designer's intent (explicitly built into its responses) or a user's attribution. Even vacuum cleaners with no designed social capabilities whatsoever may get ascribed a certain social nature by their owners based on their behaviour. In general, people ascribe social competence to robots beyond what is actually built into the robot. In a study on robots that guide an inquiry learning task in a primary school through a sequence of limited, strictly predefined and unvarying utterances, children would still spontaneously address the robot with additional questions and remarks (e.g. "hey robot, what do you think of this?") [53].

Designer's intent and user's attribution are not independent. Users attribute things to a robot, and we can actively make use of that in our designs to make the (social capabilities of the) robot more effective. Furthermore, users attribute things to a robot, so we inevitably have to deal with the socialness of a robot even if it is not their prime purpose.

Locus of initiative The final distinction that we introduce here is between robots that are meant to be autonomous, and robots that are (tele)operated. The former type of robot expresses it's "own" intent, whereas with the latter type it is the user's intent being expressed through the robot. Examples of the latter are telepresence robots (Skype-on-wheels being used to, e.g., visit school when you are ill or go to a museum when you are incapacitated, or surgery robots that allow a doctor to physically act over a distance) and the wearable body augmentation; the other types are generally meant to be more autonomous in their operation (with still some variations there: even though the medical delivery robot does generally not decide on its own which medication to deliver where, it will make its own decisions about how to act in order to fullfill that task). For the (tele)operated robots, where the locus of initiative lies with the controlling user, the question rises whom people are interacting with. The user behind the robot? Or also the robot itself? Anecdotal experience suggests that it is a mixture between the two. This could be seen to sketch a wider picture of the agency of the robot and how a user perceives its “autonomous” intelligence (not just as a piece of technology controlled by someone, but also as a perceived “selfness”) [30].

2.4 Conclusion: what is social robot technology?

This report aims to discuss the current state of technology for social robots. As can be seen from the discussions above, this will involve discussing a range of topics regarding not just their technological capabilities, but also their social and expressive capabilities. Furthermore we believe that, to do justice to the complexity of these issues, we should also incorporate the design perspective into this discussion, as well as the aims for which the robots are to be designed, since this will heavily influence how one would navigate the various dimensions discussed above. The remainder of this report will be spent on these various topics.

3 Autonomy and its basic technologies

A factory robot arm is programmed to execute a repetitive task (e.g. welding a frame or spray painting a door). Although these robots can perform these tasks automatically, usually these systems have no (or very limited) interaction nor sensing (so if the car frame would not be there, the arm would go on welding regardless). Robotic technology in the real world, and especially in the social domain, responds to events, is interactive and exhibits a level of autonomy.

In order to make that possible, a robot needs technology to **sense**, **think** and **act**, just like any other type of socially intelligent technology. Based on sensor input a robot takes a decision and acts in the real world.

Needless to say, the quality of the decisions a robot can make depends on the extent of the sensing and the complexity (elaborateness) of the decision making model. Regarding what a robot can -do- in the real world, actuation (actuators) come in many forms and requirements. An arm that can pick up a cup of coffee is completely different from an arm that has to point out something or make a gesture. For locomotion, a set of legs make a robot humaniform (and perhaps in the case of NAO very cute) but legs are not the most efficient way of travelling through a corridor in a care institute.

In many of the current experiments in research, the first two (sense and think) are replaced by a human operator. Robots that are “puppeteered” [20] (also sometimes referred to as “Wizard of Oz prototyping”) use the sensing and brain of a human operator to make the robot work. This can be done in an “honest” way (so the operator tries to act within the boundaries of the capabilities of a robot with limited sensing and reasoning capabilities) or as more speculative design (imagining really clever and smart and sentient robots).

The role of an operator or external user can also have an explicit function, for example in a case where the robot is used as telepresence device. Here the robot might be a complete transparent interface to the operator, or have a number of assistive autonomous functions, such as “drive to a group and join the conversation” as explored in the TERESA project [51].

3.1 Sensor technologies

Many of the sensors deployed in robots mimick in some way the human senses. Robots can obtain visual information through cameras [18] , auditory information through microphones [17], touch through pressure and warmth sensors [24], orientation information (balance) through Inertial Motion Unit (IMU) sensors, smell through gas sensors (although smell and taste is usually ommited), etc. These sensors may be locally attached to the robot, or a robot might make use of global sensors that are deployed in the room in which it operates (if the robot operates in a so-called smart environment). This makes some differences in what the sensors can perceive or not, and how to interpret perceptions in relation to the robot, but in principle such sensors would have roughly the same capabilities.

All these different types of sensors can be employed to deduce information about the physical state of the world [54] (where is the robot in the room; where are obstacles; how large is an object that the robot wants to pick up; how fast is the robot moving; where is sound coming from; ...) and Social Signal Processing [50] about social events and activities that are ongoing (who is in the room; who is speaking; what are they saying; what is the emotion of the user; where is the user going; what group patterns are emerging...).

There is a risk in assuming that adding a camera is enough for a robot to ‘see’ and a microphone is enough to ‘listen’. Compared to human eyes and ears, the robot equivalents are not only fundamentally different, but also require (due to their simplicity) much more processing power to generate meaningful information from the sensors. Furthermore, as will be discussed in the section about AI & intelligence, the direct information obtained from the sensors still needs to be interpreted into the appropriate higher level categories [42]. Interpretation of social behaviours and activities is heavily dependent on the underlying models and classes that the robot’s sensing modules build upon: a robot’s camera sensors can not see something that it did not already have a model for.

In the remainder of this section we will briefly go over a collection of sensing domains.

3.1.1 Propriocepsis

Many of the sensors present in a robot are necessary for proprioception - or measuring its internal state. For measuring orientation (equivalent to the human equilibrioception) robots use so-called IMU sensors. An Inertial Measuring Unit comprises of several accelerometers and gyroscopes and sometimes even magneto-compass sensors, aggregated into one orientation measurement. This can be used for example for a simple robot like Paro to detect whether it is upside-down, whether somebody is shaking it, etc. This type of sensor is also deployed in every smart-phone. It is very versatile and can be used to detect different patterns in motion, such as walking, or body activity during sleep.

Usually for every degree of freedom (everything that the robot can move, be it an arm, an eyebrow or a wheel) a sensor is needed for **feedback**. A main part of the more low-level control functions of robots deal with feedback control: moving an actuator to a certain setpoint and checking whether this setpoint is reached.

This also touches upon one of the most fundamental challenges for robotics in interaction with humans. (Feedback) control theory and application aims at getting to setpoints (i.e. a position) as fast and efficient as possible. This is of course fine for a factory robot arm, but can be outright dangerous for a robot that has to be hugged or cuddled. This means that traditional control using -only- position or velocity sensors in every degree of freedom will not be sufficient. Robots for (haptic) interaction should be able to measure forces, both excerted by the robot onto the real world, as well as the forces applied by the world (or user) to the robot.

3.1.2 World Perception

The world that the robot can perceive through proprioception is limited (but, in some cases such as a Paro robot, likely to be enough). For complex tasks such as navigation or grasping and manipulation it is necessary for the robot to build up a representation of the world around it (a so called “World Model”). This model can be quite complex, but in many cases a simple model can be sufficient as well to carry out the robot’s basic tasks. For pragmatic purposes, some robotic platforms “cheat” at this world perception by recognizing a fixed set of objects by QR codes or embedded NFC identifiers (e.g., Pleo or Aibo do this).

3.1.3 Speech

Automatic Speech Recognition is nowadays getting pretty good, but remains somewhat context and microphone dependent. Simple keyword detection can be done robustly in difficult environments; full speech understanding remains hard to do error-free so far, but developments are going quickly. Especially the existence of open source and freely accessible cloud services for speech recognition puts speech recognition in reach of applications at least where the acoustic environment is somewhat controlled.

3.1.4 Detecting People and Faces

The capability to detect where people are is generally considered crucial for social robots so the robot can show engagement and interest towards humans. Therefore, most platforms already incorporate a camera and software for at least rudimentary face detection. However, in most available systems there are still some limitations to the freedom of where the face can be located relative to the camera to be detected: the face needs to be properly in view of a camera, with not too much rotation in various directions.

3.1.5 Body Movement, Poses and Gestures

There is extensive literature on detecting poses, gestures, and other body movement from camera images [41, 40]. So far, using this in applications still requires one to use only a limited set of clearly different gestures, and movements should be relatively unobstructed in view of the camera.

3.1.6 Touch

Touch is an important interaction modality for social robots, especially given the applications in the social professions. Touch as a robot sense is done with hidden buttons or simple capacitive sensing in designated areas on the robot skin, but there are also more advanced sensing technologies being explored. Besides this “technical” perception, there is also the problem of classifying detected touches into social categories [25]. For an overview of social touch and human robot interaction, see [22].

3.1.7 Emotions and Mental States

In order to respond socially in a human context, robots should have a sense of emotion and affect. Current state of the art explores recognition of affect in multidimensional models (arousal, valence) or based on a limited set of emotional categories [44, 57]. A limitation still is that emotion classes can not always be understood in the same way across different applications, so caution is needed when transferring affect recognition models from one domain to another. It is important to also choose *task-relevant emotions* to pursue in any given application (e.g., anger, sadness, disgust, surprise, and joy are not always the most relevant emotions in any setting).

3.1.8 Final Note on On-board Versus Cloud Based Sensing

Interpretation of the information coming from a robot's sensors can cost a lot of processing power. Currently it is quite usual to offload some of that processing to cloud services; most well known is for example Google's cloud based speech processing. Of course, this has implications for privacy and safety, especially when the robot interacts with vulnerable target groups. Ethical considerations might sometimes lead to robots foregoing the more powerful cloud based solution for less powerful, but better under control on-board processing of sensor data.

3.2 Manipulation

One of the most viable attributes of robots currently employed in industry is the ability to manipulate objects and tools in physical space. Factory robots are used to pick up components and place them in an assembly or use a welding tool to assemble a car.

Key in these applications is that they work in a controlled environment where effectively little environmental awareness is necessary. A robotic arm can work using proprioception only (i.e. it has to know and control its internal position and orientation). This means that if due to a slip up in the factory a robot arm will weld in the air when the car it is supposed to work on is not there.

Many applications in care could benefit from a (mobile) robot capable of fetching objects, manipulating a bottle (or a spoon for that matter), picking up stuff from the floor, etc. Although shape of for example the design of Nao or Pepper suggest the capabilities of dexterous manipulation (i.e. they have a human like arm- and hand design), in practice these manipulators can be used only partially (keep something in their hand that has deliberately been put there) or as a tool for expression (see the next section).

Two aspects make application of robotic manipulators in a social environment particularly challenging: the need to deal with the unstructuredness of the environment and the need to deal with (unpredictable) behaviour of humans.

Much research is devoted to the engineering and construction of versatile manipulators that can work in unstructured environment [4]. The human hand is often used as benchmark since many tools and objects in social space are designed to be manipulated (as the word suggests) by a hand-like manipulator. Besides the manipulator design also the control and sensing involved is focus of current research [36].

The ability to grasp an object in space relies heavily on computer vision [18] in order to understand the shape and orientation of the object which needs to be grasped - and orientation of the manipulator with respect to this object. Although computer vision development takes rapid strides, especially through availability of open source vision toolkits like openCV [5], reliable solutions for day to day use in a home environment do not exist yet. For example the RoboCup@home challenge [23] consist of a number of (simplified) object manipulation challenges as well.

On the iCub robot research is done on grasping and manipulation [35] explicitly taking into account the interaction with a user. In one of the resulting demonstrations iCub tries to follow and grasp a red ball which a person holds in front of its face.

The robot Baxter [16] aims at bridging the gap between non-interactive factory robots (driven by precision and unaware of their surroundings) and human operators by allowing direct interaction with its manipulators by humans. During operation the arms can be touched and stopped. This is possible by stepping away from a traditional position control strategy (get to a position as precisely and efficiently as possible) but taking into account interaction with external forces – for example using impedance control [21].

3.3 Navigation

An important challenge for robots is how to get around and about. Locomotion and navigation are closely linked, and far from trivial problems, even (or especially) indoors. Since these problems are very generic (and need to be solved by also many non-social applications) the following sections follow the same structure as for example the primer by Mataric et. al. [34].

There are many styles of locomotion for robots[9], some are clearly bioinspired [39] (as in robots that move like snake, insect or quadruped). The environment social robots encounter is usually limited to an indoors / household oriented place – which means flat floors (not rough terrain). Other morphology (snakes, crawlers, drones) do not have benefits when it comes to maneuverability or efficiency indoors. Hence most of the platforms designed for use indoors use wheels.

Independent of how the robot moves from A to B, it needs guidance or control to navigate. For autonomous navigation a world model / representation is needed, combined with a method of sensing the robot’s position with respect to the real world.

3.3.1 Locomotion

Many existing robot platforms such as vacuum cleaners and telepresence robots like the Double or Giraffe use wheels. Also a humaniform robot like Pepper uses a set of wheels instead of legs. Wheels work well on flat floors, so if there are no thresholds or staircases to negotiate, a wheeled platform can work reasonably well. Depending on the exact conditions of the floor (dusty, wet, clean) and the material (wood, stone, linoleum) wheels will slip. In itself this is not a problem, however when a platform relies on proprioception to navigate (i.e. count the amount of revolutions of a wheel and use dead-reckoning to keep track of the position) it is bound to end up somewhere else then expected.

---

[23] See https://www.robocup.org/domains/3

Legged locomotion is used for many of the humaniform robots such as the Nao. Small walking robots like Nao can maintain their balance (i.e. stand in a pose and handle small disturbances) but do not do this while walking. The walking motion can be seen as replaying a static ’recording’ of walking motion. On a flat floor without disturbances this results in a motion which is not too fast (and looks even cute). A robot like Nao is however not capable of handling disturbances while walking. Small unevennesses in the floor or an external push will cause the robot to fall over. It cannot take action while falling (i.e. shift a foot or catch its fall). It will fall over. It will however also detect the fall and stand up shortly after. Needless to say, for most of the small robots designed for social interaction the legs can better be utilized (or regarded) as part of their expressiveness rather than an efficient way of getting from A to B.

3.3.2 Navigation

In many tele-operated robots especially the navigation part is solved by a human operator with a remote control. For example on a Double robot, an operator has an interface showing two camera feeds: one for the conversation (skypeon-wheels) and one camera facing downward for navigation. By arrows in the user interface the operator has to move the robot forward, backward or turn. Especially when there is some latency due to the used network (i.e. WIFI), this activity is time consuming and far from trivial. Every action (move forward) has to be observed (wait for video to be ready) and checked (did I bump in to something).

This problem with latency and network connectivity is less prominent when the robot navigates autonomously. For navigation the robot needs to know where it is with respect to a model of the world. Localisation of the robot indoors is not impossible, but requires still a substantial amount of sensors and processing power.

A common addition to measuring wheel position (odometry) is an Inertial Measurement Unit (IMU) to measure and track angles and acceleration of the platform. Still, this usually does not provide the robot with enough reliable data for navigation. Outdoor platforms can rely on GPS to give an external position reference. Indoor GPS systems are not readily available and still topic of research. In some cases existing WIFI infrastructure can be used (using received signal strength as indicator to the proximity of a known WIFI access point) but in most cases robots rely on the recognition of visual landmarks for indoor navigation [12]. Visual landmarks are detected by 2D or 3D imaging sensors such as (stereo) camera’s and LIDAR sensors.

The process in which a robot concurrently builds a world model (a map) and uses this map for navigation is called SLAM [13] (simultaneously localizing and mapping) or CML (concurrent mapping and localization). SLAM relies heavily on taking data from all available sensors and aggregating (or fusing) this data into a position estimate. This process therefore is called sensor fusion [26] and is based on incorporating a statistic likelihood of which sensors provide the best data under which conditions. This process is called Kalman Filtering [52] and uses *a priori* knowledge (a robot cannot travel infinitely fast from one side of the room to the other, when a robot is accelerating it is more likely to slip) and prioritizes data from the sensors which are presumably most accurate in the circumstances.

Although in some of the high-end products on the markets SLAM technology has been incorporated (such as some cleaner robots or security robots) the application of SLAM is still mostly limited to research platforms. It is at present a building block for development platforms like ROS [24], facilitating application for research, but because of the price of the used sensors (LIDAR, camera) and the necessary amount of processing power not widely applied in other consumer products.

4 Designing expressive behaviours

Zaga [56] describes various top-down and bottom-up approaches to designing a robot's behaviours. In general, she distinguishes: engineering and modelling methods [14, 7, 7], design methods inspired by human-human interaction theories and observation [37, 45], animation and choreography inspired methods [43, 48, 8] that make use of animation, theatric or choreography principle to guide the design of robot behaviors, embodied design methods [19, 47] which ties the design of robot behaviors to the one of embodiment of the robot combining storyboarding, physical and video prototyping, puppeteering, Wizard of Oz techniques, and experimentation. Finally, one can annotate and analyse a corpus of recordings of human intetractive behaviour in order to replicate what is seen there to arrive at proper expressive behaviour [29]. The design of expressive behaviour is typically a highly multidisciplinary field, building upon widely ranging fields such as movement science, dance, theater, communication psychology, or linguistics.

Especially in *social* robot applications, it is of crucial importance to design the right kind of expressive and conversational behaviour: a combination of motions, sounds, and possibly face expressions that communicate a certain message in the context of the interaction. Such messages are shaped to engender the illusion of emotion, action intent, or communicative intent of the robot. The exact form that this takes is closely related to the form factor of the robot. A pet-like robot will typically employ different behaviours than a humanoid robot; and a robot that is neither (fantasy shapes, or minimal expressive robots) would employ yet other forms of expressive behaviour more tailored to the affordances of that form [8, 27].

Zaga [56] also describes how Human-Robot Interaction is nowadays giving more attention to robotic technology exhibiting less human-like behaviors and roles [19, 38, 15, 33, 2]. Among other things, she discusses how this impacts the potential mismatch between expectation and real capabilities of the robot, limiting the aspiration on human-like relations between user and robot and affording different kinds of interaction. There is a shift towards robots that are interactive "things" more than artificial human replacements, e.g. in educational robotics [10, 1]; such robots exhibit behaviors afforded by their appearance rather than human like behaviours, and rely on humans' tendency to antropomorphize and to make sense of the meaning of a behavior in the interaction itself [23]. Depending on the application, the resulting behaviour may be less or more complex; e.g. Paro is mostly wriggling and squealing whereas a robot such as Zeno typically has to express complex communicative utterances that deliver specific content, so it builds more upon combinations of speech sentences with meaning-carrying gestures and face expressions.

---

[24] ROS is a widely used control software suite for (social) robots. Libraries exist for SLAm, such as openslam.org

These actual behaviours are then typically choreographed in a software platform specialised in planning and alignment of multiple partial behaviours. The exact choreography may be mostly predefined (playing canned “animations”) or planned on the fly [49]. Such a platform ideally needs to be able to seamlessly integrate top-down behaviour (deliberate planned behaviour in the context of the interaction or dialogue) with bottom-up behaviour (semi-autonomous generated behaviour in response to environmental cues, such as responsive gaze or blinking): *“This introduce a a challenge for an architecture for human robot interaction. On the one hand, the robot embodiment continuously carries out its autonomous and reactive behaviour patterns. The parameters of these may be modified on the fly based on requests by the dialogue manager. On the other hand, the dialogue manager may request deliberative behaviours that actually conflict with these autonomous behaviours, since the dialogue manager does not know the exact current state of the autonomous behaviours. The control architecture therefore contains intelligence to prioritise, balance and mix these multilevel requests before translating them to direct robot controls.”* (quote from
[11]).

5 Artificial Intelligence

Regarding the Artificial Intelligence underlying interactive robot systems, we point out two important and somewhat separate domains.

First, there are the “sense” components, that go from sensor data to interpretation. Interpretation of sensor data is often based on models that have been handcrafted, or built on training data using machine learning techniques. It is important to be aware that such components cannot understand/interpret things that the system did not yet have a model for, and a system can only understand things in a way that it fits the properties of the underlying models. For example, a robot system does not “understand a user’s emotion” in a general sense, but instead is “only” able to classify the users emotions in a limited set of classes that are incorporated in the original training data on which the emotion classifier was trained. Since this can make quite a difference (maybe the most salient emotion in a specific application was never present in the original data set!), it is important to make this background visible/understandable to both the application designer and to the end user (the professional in social care).

Second, there are the ”think” components: the decision making for the robot that decide ”what is the right thing to do now”. Such components range from limited and fixed pattern rule systems to complex systems able to learn from past cases about what is the right action in a certain situation. Also here, the distinction is not automatically visible to a user without explicitly making this clear.

Hoffmann argues strongly for “honest design” of social robots [3], in which a robot’s signalling should be always very clear about the “real” intelligence underlying the system. For example, a robot that in reality is only reading back lines from wikipedia pages that best seem to fit the user’s utterance, but presents these responses as if it is aware of the content of it’s own answers “dishonestly” suggests that it is carrying out a complex social conversation.

There is a certain inherent tension between socialness and honesty: the (perceived) socialness of a robot depends on suspension of disbelief, and involves actively designing the robot to trigger certain schemata of the user that makes them evaluate the robot as social. This may help the robot carry out its task more effectively. but actively designing a kind of fake illusion of social nature in the robot, which runs counter to the honest design. This may lead to the robot overpromising capabilities, the user overattributing qualities to the robot, and a general mystification of robot capabilities. This particular tension has not yet been resolved, and is a current focus point for some researchers in Human-Robot Interaction design.

6 Conclusions

When looking at the scoping report on Statistics and Trends in Social Robots in Social Care, delivered together with this report as part of the PRoSPERo project, one can see there is a multitude of existing social robots in the social professions. In social care there are also great opportunities for these robots to make a meaningful difference.

The materials in this report set out a number of topics which could play a role in the envisioned curriculum (or: Module Descriptors) that PRoSPERo aims to deliver as ultimate outcome. We argue that it remains important to keep a nuanced and "de-mystified" view on social robots, both regarding what a social robot is, and regarding what are possible useful and worthwhile contributions of a robot in an application context. Although social care professionals should probably not be educated into a deep and generative understanding of the technology behind social robots, they should be aware at least to some extent of the possibilities and impossibilities of the state of the art technology embedded in the kind of robots they are likely to encounter in their work. Furthermore they should be aware of the challenges and aims in the *design work* in social robotics, especially regarding form and behaviour. We expect, on the basis of our work in tinkering and design workshops for non-technologists, that a certain hands on experience in these topics goes a long way to addressing this awareness. Finally, when looking at the example applications discussed in the Statistics and Trends report, we argue that there is an aspect of applying social robots that is very much underrepresented in the literature: the design of new activities with existing social robots, as opposed to

The current image of social robots as projected through popular media is not really helpful in light of our educational and pedagogical aims; we think a big step in addressing that problem is a combination of expectation management and developing good teaching materials to help social care professionals confidently navigate the field of social robotics.

References

[1] Ackermann, E. K. Playthings that do things: A young kid's "incredibles"! In *Proceedings of the 2005 Conference on Interaction Design and Children* (New York, NY, USA, 2005), IDC '05, ACM, pp. 1–8.

[2] Alves-Oliveira, P., Arriaga, P., Paiva, A., and Hoffman, G.

YOLO, a robot for creativity: A Co-Design study with children. In *Proceedings of the 2017 Conference on Interaction Design and Children* (June 2017), ACM, pp. 423–429.

[3] Anderson-Bashan, L., Megidish, B., Erel, H., Wald, I., Grishko, A., Hoffman, G., and Zuckerman, O. The greeting machine: An abstract robotic object for opening encounters.

[4] Bicchi, A. Hands for dexterous manipulation and robust grasping: A difficult road toward simplicity. *IEEE Transactions on robotics and automation 16*, 6 (2000), 652–662.

[5] Bradski, G., and Kaehler, A. *Learning OpenCV: Computer vision with the OpenCV library*. ” O’Reilly Media, Inc.”, 2008.

[6] Breazeal, C., Dautenhahn, K., and Kanda, T. Social robotics, 2016.

[7] Breazeal, C. L. *Designing sociable robots*. MIT press, 2004.

[8] Bucci, P., Cang, X. L., Valair, A., Marino, D., Tseng, L., Jung, M., Rantala, J., Schneider, O. S., and MacLean, K. E. Sketching cuddlebits: Coupled prototyping of body and behaviour for an affective robot pet. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), CHI ’17, ACM, pp. 3681– 3692.

[9] Choset, H. M., Hutchinson, S., Lynch, K. M., Kantor, G., Burgard, W., Kavraki, L. E., and Thrun, S. *Principles of robot motion: theory, algorithms, and implementation*. MIT press, 2005.

[10] Cila, N., Smit, I., Giaccardi, E., and Krose, B.¨ Products as agents: Metaphors for designing the products of the IoT age. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (May 2017), ACM, pp. 448–459.

[11] Davison, D., Gorer, B., Kolkmeier, J., Linssen, J., Schadenberg, B., van de Vijver, B., Campbell, N., Dertien, E., and Reidsma, D. Things that make robots go hmmm : Heterogeneous multilevel multimodal mixing to realise fluent, multiparty, human-robot interaction. In *Proceedings of eNTERFACE ’16* (2017), K. Truong and D. Reidsma, Eds., Telematica Instituut / CTIT, pp. 6–20.

[12] DeSouza, G. N., and Kak, A. C. Vision for mobile robot navigation: A survey. *IEEE transactions on pattern analysis and machine intelligence 24*, 2 (2002), 237–267.

[13] Dissanayake, M. G., Newman, P., Clark, S., Durrant-Whyte, H. F., and Csorba, M. A solution to the simultaneous localization and map building (slam) problem. *IEEE Transactions on robotics and automation 17*, 3 (2001), 229–241.

[14] Dragan, A. D., Lee, K. C., and Srinivasa, S. S. Legibility and predictability of robot motion. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction* (Piscataway, NJ, USA, 2013), HRI ’13, IEEE Press, pp. 301–308.

[15] Fink, J., Lemaignan, S., Dillenbourg, P., Retornaz, P., Vaus-´ sard, F., Berthoud, A., Mondada, F., Wille, F., and Franinovic,´
K. Which robot behavior can motivate children to tidy up their toys?: design and evaluation of ranger. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (Mar. 2014), ACM, pp. 439–446.

[16] Fitzgerald, C. Developing baxter. In *2013 IEEE Conference on Technologies for Practical Robot Applications (TePRA)* (2013), IEEE, pp. 1–6.

[17] Gomez, R., Kawahara, T., Nakamura, K., and Nakadai, K. Multiparty human-robot interaction with distant-talking speech recognition. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY, USA, 2012), HRI ’12, ACM, pp. 439–446.

[18] Hartley, R., and Zisserman, A. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[19] Hoffman, G., and Ju, W. Designing robots with movement in mind. *Journal of Human-Robot Interaction 3*, 1 (Mar. 2014), 89–122.

[20] Hoffman, G., Kubat, R., and Breazeal, C. A hybrid control system for puppeteering a live robotic stage actor. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication* (2008), IEEE, pp. 354–359.

[21] Hogan, N. Impedance control: An approach to manipulation. In *1984 American control conference* (1984), IEEE, pp. 304–313.

[22] Jung, M. *Socially intelligent robots that understand and respond to human touch*. PhD thesis, University of Twente, Netherlands, 6 2017. CTIT Ph.D. thesis series no. 17-437 SIKS dissertation series no. 2017-26.

[23] Jung, M. F. Affective grounding in Human-Robot interaction. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (Mar. 2017), ACM, pp. 263–273.

[24] Jung, M. M., Cang, X. L., Poel, M., and MacLean, K. E. Touch challenge ’15: Recognizing social touch gestures. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (New York, NY, USA, 2015), ICMI ’15, ACM, pp. 387–390.

[25] Jung, M. M., Cang, X. L., Poel, M., and MacLean, K. E. Touch challenge ’15: Recognizing social touch gestures. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (New York, NY, USA, 2015), ICMI ’15, ACM, pp. 387–390.

[26] Kam, M., Zhu, X., and Kalata, P. Sensor fusion for mobile robot navigation. *Proceedings of the IEEE 85*, 1 (1997), 108–119.

[27] Karreman, D. *Beyond R2D2 - The design of nonverbal interaction behavior optimized for robot-specific morphologies*. PhD thesis, University of Twente, Netherlands, 9 2016. SIKS dissertation series no. 2016-36 ; CTIT Ph.D. Thesis Series, ISSN: 1381-3617, No. 16-404.

[28] Kennedy, J., Baxter, P., and Belpaeme, T. Comparing robot embodiments in a guided discovery learning interaction with children. *International Journal of Social Robotics 7*, 2 (2015), 293–308.

[29] Kipp, M. *Gesture generation by imitation : from human behavior to computer character animation*. PhD thesis, 2003.

[30] Lee, M. K., and Takayama, L. ”now, i have a body”: Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2011), CHI ’11, ACM, pp. 33–42.

[31] Leyzberg, D., Spaulding, S., Toneva, M., and Scassellati, B. The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the annual meeting of the cognitive science society* (2012), vol. 34.

[32] Li, J. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies 77* (2015), 23—-37.

[33] Lupetti, M. L. Shybo. an open-source low-anthropomorphic robot for children. *HardwareX 2*, Supplement C (Oct. 2017), 50–60.

[34] Mataric, M. J., and Arkin, R. C.´ *The robotics primer*. Mit Press, 2007.

[35] Metta, G., Natale, L., Nori, F., Sandini, G., Vernon, D., Fadiga, L., Von Hofsten, C., Rosander, K., Lopes, M., Santos-Victor, J., et al. The icub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks 23*, 8-9 (2010), 1125–1134.

[36] Murray, R. M. *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[37] Mutlu, B., Forlizzi, J., and Hodgins, J. A storytelling robot: Modeling and evaluation of human-like gaze behavior. In *Humanoid robots, 2006 6th IEEE-RAS international conference on* (2006), Citeseer, pp. 518–523.

[38] Ozg¨ ur, A., Lemaignan, S., Johal, W., Beltran, M., Briod, M.,¨ Pereyre, L., Mondada, F., and Dillenbourg, P. Cellulo: Versatile handheld robots for education.

In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI* (2017), IEEE, pp. 119–127.

[39] Pfeifer, R., Lungarella, M., and Iida, F. Self-organization, embodiment, and biologically inspired robotics. *science 318*, 5853 (2007), 1088–1093.

[40] Poppe, R. Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding 108*, 1 (2007), 4 – 18. Special Issue on Vision for Human-Computer Interaction.

[41] Poppe, R. A survey on vision-based human action recognition. *Image and Vision Computing 28*, 6 (2010), 976 – 990.

[42] Reidsma, D., op den Akker, R., Rienks, R., Poppe, R., Nijholt, A., Heylen, D., and Zwiers, J. Virtual meeting rooms: from observation to simulation. *AI & SOCIETY 22*, 2 (Nov 2007), 133–144.

[43] Saerbeck, M., Schut, T., Bartneck, C., and Janse, M. D. Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2010), CHI ’10, ACM, pp. 1613–1622.

[44] Sariyanidi, E., Gunes, H., and Cavallaro, A. Automatic analysis of facial affect: A survey of registration, representation, and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence 37* (2015), 1113–1133.

[45] Sauppe, A., and Mutlu, B.´ Design patterns for exploring and prototyping human-robot interactions. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2014), CHI ’14, ACM, pp. 1439–1448.

[46] Share, P., and Pender, J. Preparing for a robot future? social professions, social robotics and the challenges ahead. *Irish Journal of Applied Social Studies 18* (2018).

[47] Sirkin, D., Mok, B., Yang, S., and Ju, W. Mechanical ottoman: How robotic furniture offers and withdraws support. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (Mar. 2015), ACM, pp. 11–18.

[48] Szafir, D., Mutlu, B., and Fong, T. Communication of intent in assistive free flyers. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (Mar. 2014), ACM, pp. 358–365.

[49] van Welbergen, H., Reidsma, D., and Zwiers, J. Multimodal plan representation for adaptable bml scheduling. *Autonomous agents and multi-agent systems 27*, 2 (9 2013), 305–327. eemcs-eprint-22871.

[50] Vinciarelli, A., Pantic, M., and Bourlard, H. Social signal processing: Survey of an emerging domain. *Image and Vision Computing 27*, 12 (2009), 1743 – 1759. Visual and multimodal analysis of human spontaneous behaviour:.

[51] Vroon, J. Responsive social positioning behaviors for semi-autonomous telepresence robots. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY, USA, 2017), HRI '17, ACM, pp. 383–384.

[52] Welch, G., Bishop, G., et al. An introduction to the kalman filter.

[53] Wijnen, F. M., Davison, D. P., Reidsma, D., Van Der Meij, J., Charisi, V., and Evers, V. Now we're talking: Learning by explaining your reasoning to a social robot. *Submitted for review*.

[54] Yan, H., Ang, M. H., and Poo, A. N. A survey on perception methods for human–robot interaction in social robots. *International Journal of Social Robotics 6*, 1 (Jan 2014), 85–119.

[55] Zaga, C. Something in the way it moves and beeps: Exploring minimal nonverbal robot behavior for Child-Robot interaction. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on HumanRobot Interaction* (Mar. 2017), ACM, pp. 387–388.

[56] Zaga, C. *The design of robothings for collaborative play*. PhD thesis, University of Twente, in preparation.

[57] Zeng, Z., Pantic, M., Roisman, G., and Huang, T. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE transactions on pattern analysis and machine intelligence 31*, 1 (1 2009), 39–58. 10.1109/TPAMI.2008.52.