**Developing Socially-Inspired Robotics through the Application of Human Analogy: Capabilities and Social Practice**

Neil McBride

Centre for Computing and Social Responsibility

De Montfort University

The Gateway

Leicester

LE1 9BH

United Kingdom

nkm@dmu.ac.uk tel: 0116 2078500

orcid: 0000-0002-7110-0709

**Abstract**

Socially-inspired robotics involves drawing on the observation and study of human social interactions in order to apply them to the design of sociable robots. As there is increasing expectation that robots may participate in social care and provide some relief for the increasing shortage of human care workers, social interaction with robots becomes of increasing importance. This paper demonstrates the potential of socially-inspired robotics through the exploration of a case study of the interaction of a partially-sighted social worker with a support worker. This is framed within the Capability Approach in which the interaction of a human and a sociable robot is understood as resulting in a collaborative capability which is grounded the relationship between the human and the robot rather than the autonomous capabilities of the robot. The implications of applying the case study as an analogy for human-robot interaction are expressed through a discussion of capabilities and social practice and policy. The study is attenuated by a discussion of the technical limits of robots and the extensive complexity of the social context in which it is envisaged sociable robots may be employed.

**Keywords**

Sociable robots, socially inspired robotics, trusting robots, capability approach.

1. **Introduction**

In the last decade, technical development in humanoid robots have triggered immense research on the use of robots in society: in social situations where there is both emotional and physical interactions with humans in complex and dynamic social environments. Robots are likely to be employed in social care (Houses of Parliament, 2018). Robots may interact with humans in business services situations, such as university reception (Nisimura et al, 2002) and in shopping centres (Chen et al, 2017).

The UK government envisions the development of health care robots which will contribute to the grand challenge of an aging population and the goal of giving people five years of longer healthier life (UK Government, 2019). Using care robots “*could help provide the UK’s dedicated adult social care sector with more assistance for those who need it most*.” The government aim to “*develop robots to one day fulfil tasks such as helping an elderly person up after a fall and raising the alarm, delivering food to an older person at mealtimes, and even ensuring they take crucial medication at the correct time.*” These robots will have “*a clear set of rules in order for them make effective decisions*” and “*demonstrate principles like respect, fairness and equality enabling them to eventually be used in environments like care homes and hospitals*.” They will “*work in tandem with professionals to assist and complement their work, and help relieve pressures*.”

Such visions of the role of robots in society pitches the robotics away from excavating on Mars, packing boxes of sweets or painting Airbus A350s into the care home and the hospital, lifting the elderly out of bed and changing nappies (Hall, 2019). This functional shift will require that robots interact socially on a daily basis, even if their basic domestic tasks are primarily repetitive and physical. This expanding expectation of social interaction and connection with robots creates two problems. Firstly, there is the problem of gaining a social understanding of the gap between human socialisation and the limited rules-based or pattern-based approaches which AI offers. Secondly there is the problem of how reasonable design of social scripts for robots can be achieved. We need to avoid the “cargo cult science” that Collins attributes to sociable machines such as Cog and Kismet (Collins, p138n).

In engaging robots in society in a less constrained and hence more social and uncertain way, we must tackle the disjuncture between human expectations and the capabilities of technological artefacts such as sociable robots (Suchman, 2007). This problem takes on a greater importance because the social shift of robots outside of the constraints of manufacturing and into the public sphere of social care. The prioritising human interactions represent a shift from a product-based functioning to a service-based functioning. In a service-based activity, quality is represented by the gap between expectation of a service and perception of that service, neither of which are solely objective evaluations. Rather than just managing the quality of the production process, we must now manage the expectations and perceptions of the user who exercises the service in partnership with the producers. Thus the focus of social interaction becomes relational.

This means that we cannot solely design the actions of the robot, but we must design the social interaction of the robot with the user. This involves understanding and scripting both user and robot activities and the relationship between them. Thus understanding and managing the role of robots within society takes on a social context which was not of so great significance in manufacturing roles or guiding robots on Mars. Our study and design of interaction of robots in everyday society and life must go beyond planning motions and tasks and goal achievement to understanding the complex personal and social interactions between robots and individuals within a complex social context.

This paper establishes the concept of socially-inspired robotics in which experience and understanding of human social interaction is used as analogies for designing and constraining human-robot social interaction. Socially-inspired robotics involves drawing on examples of human social interaction as a basis for anticipating and planning social interaction between humans and robots.

Social interaction with robots involves the compensating for degraded human capability and the development of new capabilities. The capabilities thus developed are joint capabilities between the robot and the human. In understanding the role of socially-inspired robotics I suggest that Sen’s capability approach (Alkire, 2005) provides a useful tool for developing socially-inspired robotics.

I will firstly define the idea of socially-inspired robotics and illustrate its relation to biologically-inspired robotics. This raises concerns about the feasibility of such an approach and its limitations which I will discuss. I will illustrate the approach by drawing on a narrative from the experienced interaction between a social worker and her support worker. Through this narrative example, a wide range of artefacts as an outcome of a design process can be envisaged which include policies, service-level agreements and manufacturer’s instructions as well as socially-inspired algorithms. These can be understood both in the individual robot supported capabilities and in the wider social context.

1. **Socially Inspired Robotics**

The future potential for robots to work with humans in social situations in the field depends partly on advances in robot engineering. Robots can reflect human emotion (Fellous, 20043), and detect and respond to emotional signs (for example, the EmpathyBot; Dexter, 2019). They may interact in a domestic situation, by for example, taking foods to the owner (New Atlas 2019). Robot human interactions are envisaged in medical situations where the robot may be required to physically intervene (McMullan, 2016) or support the patient by administering medication (Anderson and Anderson, 2010). These socially interactive robots can be distinguished from robots used in teleoperation such as the mars rover (Vertesi, 2012). They use a range of social cues such as gaze and gesture; they establish social relationships and respond to a wide range of human social behaviour (Fong et al, 2003). The development of sociable robots focuses on peer-to-peer human-robot interaction. This involves programming emotions, dialogue, natural embodied cues, personality and character. Algorithms must model, predict and drive social interaction between the robot and the human.

How we design robot human partnerships will influence social culture change and social policy concerning the use of robots in general society. Given the importance of human interaction in social robotics, the issue of how much control and computation should be allocated to the robot is of key importance. Shared autonomy is important in any human – human relationship and should be considered as key to effective robot human relationships. Everyday robots, in close contact with people will have to work with sensitivity not just to the social context, but also to the division of tasks, differences in skills and understanding, perception concerning the state of the human robot partnership and implementation of appropriate density of communication (McBride and Hoffman, 2016). The robot must perform socially. Behaviour and function must be customised to social and cultural context (Knight, 2014).

In order to develop the social role of robots it will be important to apply insights from human social relations and behaviour and develop socially-inspired robots in the same way as biology has been explored for biological metaphors which can be used in computing and the development of robotics. Biologically inspired robotics has affected the development of robots in significant ways. The morphology of robots has been inspired, for example, by ants (Garnier, 2011), snakes (Ye et al, 2003).and dragonflies (Grifantini, 2008) . Approaches to locomotion have been modelled on dogs (Kosaka, 2013), spiders (Pavic, 2010) and the upright gait of humans and primates (Byagowi and Kopacek, 2010).

In the same way that mechatronic strategies have been developed from biological models, the social programming of robots, and their preparation to coexist with humans in social situations must be informed by an understanding of social situations and turning to social examples as analogies from which to draw inspiration in the development of social algorithms for robots. Such analogies need to be drawn from a wide range of social situations where human individuals and groups interact. They must provide examples of the explicit and tacit interactions which lubricate human society. Many will concern individual interactions including cooperation to achieve tasks, supportive interaction in a social situation such as a hospital or a school and communication to establish relationships.

Socially inspired robotics involves drawing on behavioural patterns expressed in human interactions as a basis for designing robot behaviour. Both group and one-to-one individual behaviour can be analysed and used to develop algorithms which enable socially-appropriate robot behaviour. Socially inspired robotics will support the development of sociable robots which engage in social interactions, learn from their own experience and communicate both with other robots and with humans. Socially-inspired approaches have been used to develop path planning algorithms to help robots move through a crowd (Muller et al, 2008) and in developing strategies for robots to approach people (Ramirez et al. 2016). In the case of Chen et al’s Kejia shopping assistant (Chen et al, 2017), while to some extent socially-inspired, development concentrated on navigating through a large mapped environment, reaction to crowd situations, detecting humans and recognising commands, which is a long way from the social interaction imagined by Stahl et al (2014) for an empathic care robot.

Breazeal (2002, p27) drew inspiration from developmental psychology and explored early infant-caregiver interactions as a basis for developing Kismet, a social robot. Gleeson et al (2013) sought to understand gestures in human-human interaction in the context of simulated assembly tasks. Once identified and classified, such gestures could be developed for robots such that the interaction between humans could be used in human-robot interactions. However, such applications only touch on the conversation analysis and emotional understanding which may be expected from sociable robots, although conceptually and practically difficult to deliver.

This paper offers an illustration of how socially-inspired robotics may work. A real-life social relationship described. From this some social interaction principles and themes are developed. These may then be used as the basis for developing social algorithms and for judging the efficacy of robot behaviour in a social environment.

1. **Social Context of Socially-Inspired Robotics**

The shift from designing and implementing functions to designing social interactions, which the approach of socially-inspired robotics is directed at, requires a move away from structure and function,. Suchman et al (2002), in considering the importance of the prototype, recognise that the design and release of technology into society is an ongoing process. This process needs to address the practice of assembly, demonstration and performance. This, Suchman et al (2002) suggest, requires considering the social configuration of a technology across a diverse milieu of locations and situations.

However, the complexity of the human social environment compared with a physical environment, is such that that it may be beyond our engineering capabilities to match the variety of interactions within the social environment. The demands placed on a sociable robot in navigating a populated space, sensing and responding to non-verbal signals and engaging in verbal interaction are substantial. So any sociable robot must be constrained to a social environment in which the variety of social response the robot is capable of matches the variety of social interactions. One goal of socially inspired robotics would then be to find the boundary beyond which the social capabilities of the robot falls short of the social variety required.

The use of socially-inspired robots may help expose the difficulties and limitation of human-robot interaction. Attempts to apply biologically-inspired robotics in, for example, mimicking human walking reveal the complexities of getting a robot to even go up some stairs without falling over. There may be a significant gap between the expectation expressed, for example, in the UK Government’s press release and the actual social capabilities. There is uncertainty concerning capturing, identifying and interpreting emotional signals in computer systems (Landowska, 2019). Robots cannot achieve the social effect of a sympathetic touch simply by simulating the human touching action (Willemse et al, 2017). Face and speech recognition and sound location are fraught with difficulties, particularly in a noisy social environment (Deniz et al, 2007). Natural language generation is data-intensive and requires sufficient training data (Foster, 2019). These and other substantial limitations should be considered in a social context.

The difficulties in developing sociable robots are not only located internally but also externally in the complex social and economic environment which sociable robots would be expected to interact in. Robots require significant support in themselves. Robots in DARPA challenges are accompanied by teams of engineers ready to recover lost robots. The adequate use of sociable robots requires understanding on how to communicate with tem and get the best out of them. Suchman (2006), on visiting the MIT AI lab, found that Cog was not actively being worked on, as a consequence of which its software code was degrading due to lack of maintenance and upgrading. Also, it was difficult to elicit intelligible behaviour from Kismet due to a lack of training in the reading of Kismet’s action and how to respond (Suchman, 2006). My experience of a Pepper robot at a Eurostar station was that, despite reading the public interaction instructions, I could not elicit a response beyond its head moving to follow my movements.

The knowledge required by a sociable robot extends beyond practical knowledge concerning navigating physical barriers, but also social knowledge: abilities in listening, awareness, and social engagement that many humans struggle with. The problem is that any attempt to reduce these to rules, as a Asperger adult will know, is fraught with difficulties. The social interaction of passing someone on a sidewalk would require a large set of rules and a means of determining the context and coping with its changing (Collins, 2010, p130).

In social interactions, whether, for example teaching an autistic child or companionship for an elderly patient, the nature of the human / robot interaction fundamentally changes from one that concerns embodiment and collaboration to complete physical tasks to one that concerns social interaction and communication. This represents a shift from practical knowledge concerning physically mediated problem solving such as a robot hand solving Rubik’s cube to social and culture knowledge gained through socialisation within a practice. Robots may be able to handle the somatic tacit knowledge required to handle actions such as the twist and stretch action in bread making (Collins, 2010,p142) but social interaction requires a more sophisticated tacit knowledge, the collective tacit knowledge which drives social life (Collins,2010, p117). There is a significant difference between riding a bike to negotiating through city traffic which incorporates social context (Collins, 2010, p121).

Our social interaction also references embodiment. As such the physical embodiment of a sociable robot is an additional design issue (Deng et al, 2019). Anthropomorphism is positively correlated with perceived intelligence and the perception of autonomy (Barlas, 2019). Such perception of autonomous decision making on the part of the robot and perception of reduced agency on the part of the human can be manipulated through verbal description of the extent of robot autonomy. This raises ethical questions concerning the extent to which the designers of the robot are deceptive about the robot’s social potential, and the robot behaviour appearing to be more socially sensitive than it actually is (Danaher,2020). Robots can ‘push our Darwinian buttons’ tricking us into a fantasy of reciprocation (Turkle,2010).

Such issues concerning the computational capabilities of the robot, the complex social environment, whether work-based or domestic in which it will be required to act, and the network of industries and engineers required to support it should significantly dampen our expectations for sociable robots. However, if we constrain robots within limited social situations and consider the collaborative capability at the heart of the human robot relationship as the prime design issue, rather than the internal autonomy of the robot, it may be possible to develop socially-inspired robots which support human flourishing.

1. **Research Approach**

The approach I have taken to develop this work on socially-inspired robotics is inspired by Denzin’s approach to interpretive interactionism (Denzin, 2001). This approach recognises the importance of the experience of the researcher. In any research we bring our worldview and our personal and professional experience to bear. Our approach to and the outcomes we generate from our research are born out of who we are and how we live. My research process can be characterised in four steps: reflection, epiphany, ordering and transformation.

Reflection. The researcher is in a continuous process of reflection, operating consciously and sub-consciously. I have worked on robot ethics, and had reflected on the ethical gap between AI and humans. Attempts to represent ethics as rules encoded in algorithms in the robot may not result in an ‘ethical robot’. As suggested in BBC Science Focus Magazine (Science Focus 2020) robots won’t think like us. In enabling the robot to operate ethically we have to constrain the environment so that there are few ethical options. The focus must be on the human – robot interaction, and on the exchange of information, not on a set of limited ethical rules encoded into the robot. My reflections extended to considering sociable robots and the expectation of social interaction with a robot. Additionally, I was drawn to the possibility of applying the capability approach, which I had used in development studies, to my understanding of sociable robots.

Epiphany. Such reflection never ceases and continues through initial work, drafts, revisions, and responses to referees and will no doubt continue. But at some point there is a realisation of a way forward which determines the direction of the work. This is the epiphany or awakening. This draws on our lived experience. It occurred to me that the relationship between a partially sighted social work and a close relative had parallels with the relationship between a human and a care robot. Not that I considered the relative could be replaced by a robot, but rather that considering the relationship might provide direction on how to develop a more sociable robot, and how to sensitise robot engineers and the public to the possibilities and the limitations of robot-human interaction. Furthermore, I had done some study on biologically-inspired computing, a field which has developed through areas such as genetic algorithms and neural computing, to the physical mimicking of the animal world in robots and wondered if a similar approach could be used for sociable robotics.

Ordering. I had observed and discussed my relative’s work and the progression of the relationship over a period of two years. My ordering of this experience required the development of a succinct narrative which identified key events and changes in the relationship between the partially sighted worker and the support worker. The narrative had to be structured and succinct. Through the work of ordering I developed an analysis of the key aspects of that relationship and its progression (See Table 1).

Transformation. While there was no temporal boundary between ordering and transformation, my focus moved from describing the situation of the support worker to applying it as an analogy to sociable robot. The transformation involves a move from study and description of the human interaction to the locating of that interaction between human and robot. It became clear that there is a lot of responsibility remaining on the human side of the human-robot relationship, that trust is easily lost, that roles and tasks must be clearly defined, and that the relationship is dynamic and changing and can easily degrade as too much is expected of the robot. For the purpose of this paper I concentrated on a capability approach to sociable robot collaboration with humans and the development of social policy. I note that such transformation could go further, for example, into the development of human–robot interaction scripts.

1. **Social Capabilities**

In exploring socially inspired robotics we cannot rely on material equivalents but must look for social equivalents. Our focus moves from physical capabilities such as walking and lifting to social capabilities such as proximity management and intervention. Social capabilities will build on physical capabilities. Contributing to a social group requires sight, hearing and mobility on which is built the social capability. The capability approach provides a way to evaluate an example of social behaviour in order to identify the capability which can be attributed to the robot and the capability required in relation to the robot. The role of a sociable robot may be to amplify an existing human capability or compensate for an absent or compromised human capability. This will require human-robot cooperation such that the robot is cognisant of the limits of the human and neither encroaches on the human by over-involvement nor fails to support the human by too little intervention.

The capability approach revolves around what people can actually do. Their freedom is located in their capabilities. Being capable of something leads to a functioning in which the capability is applied to achieve a goal and hence improve the quality of life. A classic illustration of the capability approach involves considering the riding of a bicycle. A woman rides a bicycle to town in order to sell her cassava. She is capable of riding a bicycle because she has a bicycle, which requires the funds to buy it and maintain it. She has the skill of riding it. There is sufficient road between the village and town to ride the bicycle. She is protected from attacks on the way. Her husband doesn’t stop her from riding into town. Having a bicycle is not enough. The capability is complex and socially located. The woman’s capability of riding a bicycle to town enables a functioning– the selling of the cassavas at a profit without a middle man taking the profit. This gives her the capability to send her son to school because she can afford the fees. However, the paying of the fees requires an additional capability to transfer the money to the school via a mobile payment app.

Whereas the bicycle is a passive object in terms of the development of the capability, a robot is active and has capabilities of its own. Consider a blind person, moving in a crowded room during a party accompanied by a robot that can see, identify people, walk to the right guest and introduce the guest to the blind person. The robot has capabilities independent of the human. But these capabilities, combined with those of the human, .produce a collaborative social capability which enables the human to function in a social situation. While in this example, the robot contributes physical capabilities to enable the social capability, in another situation, a robot interpreting faces could provide emotional information to an autistic person in a social situation. The pairing of the human and the robot capabilities fits the human for a role in society which is viable and acceptable. For example, the pairing of a robot and an elderly patient may result in the capability of social affiliation (Coeckelbergh, 2008).

The expression of a collaborative capability requires substantial communication between the robot and the human. This is illustrated in the case study. The human (represented by the social worker) must acknowledge both the capability of the robot (represented by the support worker) and the limitations of her own capabilities. Consequently she neither attempts actions for which the robot is programmed to compensate for her suboptimal capability, nor assumes that the robot has her own capabilities, which would set an unsustainable expectation of the robot. Similarly the robot should be programmed only with the capabilities which compensate for the human’s inadequacy. The robot will also require information about the human’s capabilities in order to avoid inappropriate intervention which might disempower the human.

Emergent capabilities require socially-inspired design which is cognisant of both the robot functional requirements and the human activity. This involves an aligning or matching of the two to achieve the desired end. It is necessary to understand how the human and the robot activity act together to achieve human flourishing in a social context. Developing the emergent capability also requires the design of social communication. It is the communication between the robot and the human which will enable the generation of the capability.

In the next section I present the story of Mary and the social worker. As you read, reflect on what collaborative capabilities are developed as a result of the interaction. How do the capabilities relate to the functioning of the social worker? What is the role of trust? How does that trust break down and what are the consequence of that breakdown? How is trust restored? One result of study of the social interaction presented here is that socially-inspired design issues can be identified. Furthermore, necessary social policy for the development and deployment of sociable robots can be suggested based on the analogy of Mary and the social worker.

1. **The Human Support Worker as an analogy for the human robot relationship**

Mary was a support worker for a partially sighted social worker. She sat next to the social worker and took notes for her, found phone numbers, and completed case notes. Mary helped the social worker in meetings with clients. She drove the social worker to the meetings. When walking to a meeting, she walked next to the social worker, intervening if there was a major obstacle, otherwise walking shoulder to shoulder, so that the social worker could get guidance in a discrete manner. In meetings with clients, Mary helped by making sure records that have to be signed are accessible, indicating where there are obstacles and, where appropriate, communicating to the client.

For Mary, the nature of the role was established at the job interview. The job advertisement suggested what skills were required, and through the interview a judgement was made of that Mary had the skills, and just as importantly, that Mary would fit into the culture of the social work team. Mary was introduced to the social worker, and they settled down to work. In a probationary period, the social worker could further observe how Mary worked and whether the relationship would work in the long term.

Mary needed to empathise so that she knew when help was needed and when it wasn’t. She depended on cues from the social worker in order to carry out her support job. She needed to practice patience and develop understanding. Equally the social worker could not complete some tasks without Mary’s help and struggled with visits if Mary was not next to her. Hence the relationship was one of interdependence which required mutual understanding of each other’s abilities and limitations. It required joint understanding of the goals and tasks and willingness to receive advice and change actions and behaviour.

The social worker may have wished Mary to work faster, but acted patiently, aware that Mary was still learning about the administrative systems. Mary respected the social worker’s way of working, her use of English, her habits and did not attempt to intervene. As Mary learnt about the job, she became more competent. And as the social worker’s confidence in Mary grew, so her expectations of Mary grew.

While Mary’s boss was ultimately responsible for all tasks and outcomes, Mary had her own responsibilities, including being on time, accuracy in filling out case notes and care in protecting the privacy of clients. If Mary started telling the social worker what decisions she should make or criticised a care plan for a service user, this would be stepping outside her role and crossing the boundary. Equally, if the social worker expected Mary to start advising her and to make decisions about service users she would not be respecting Mary’s role, its limitation and boundary. The goal of the relationship would then be compromised; the task might not be completed.

At some point trust broke down. The social worker became impatient with Mary. The social worker’s perception was that Mary wasn’t finding details fast enough; she was making too many mistakes on the administrative computer system; she was not listening to what the social worker was saying. And Mary felt the social worker was not making any allowances for the limitations Mary had. Neither did the social worker understand the difficulties Mary had with the computer system because, being partially sighted, the social worker had never used the system. The social worker’s impatience and Mary’s anxiety grew. In this breakdown of trust, the social worker kept trying to check what Mary had done when Mary was quite competent; she reprimanded Mary for not remembering a detail which it was not her responsibility to remember, or for not doing something she actually had done. Minor errors by Mary became a source of irritation instead of being tolerated. Expectations needed to be renegotiated. The social worker needed to recognise the extent of Mary’s capabilities and responsibility.

Such loss of trust created uncertainty about Mary’s role, increased the number of errors and slowed the completion of tasks down. The restoration of trust in the relationship required both a renewal and revision of knowledge about the relationship, the task and the boundaries, and the exercise of the virtues of honesty and patience. As the social worker made a conscious decision to be more patient and to accept certain mistakes, the relationship was restored. Mary’s anxiety decreased and her confidence grew.

Mary accompanied the social worker on many visits. Sometimes, when walking she had to take control, grabbing the social worker’s arm to stop her walking into a lamppost, or stopping her walking into traffic. In these cases communication was increased and Mary explained what had happened.

Table 1 identifies specific issues for the social worker and the support worker.

|  |  |
| --- | --- |
| PARTIALLY SIGHTED WORKER | SUPPORT WORKER |
| Proximity | Positioning relative to partially sighted worker. Sitting next to the worker, discrete intervention. |
| Varying means of interaction. Talk, discrete contact.  | Varied range of tasks. Note taking, finding information, typing in information, walking, driving. |
| Awareness of standard codes and policies | Awareness of standards and policies.Data protection policy awarenessTechnical terms, languages and systems. |
| Understanding of goal and purpose. | Understanding of goal and purpose. |
| Agreed role of support worker  | Role, tasks, capabilities defined. Probationary period to adjust human behaviour to get the right fit. |
| Transparency understanding of limitations | Respecting social worker’s way of working.Creating predictability – through codes, boundaries, task definition etc. |
| Tacit and explicit knowledge. | Requires ongoing dynamic knowledge of what each partner is doing.Understanding tacit expressions of wishes. |
| Attenuation of behaviour: Partially sighted worker adjusts expectations and behaviour to the capabilities of the support worker.  | Support worker judges when to intervene. But follows instructions of boss carefully. Letting alone as important as intervening.  |
| Understanding and declaring assumptions.Assumptions about what the support worker’s role is and where the boundaries are must be surfaced and declared. | Limited responsibility of within the context of the relationship with the human worker, the confines of the social environment, and the rules, codes, norms within which one is working. Not assuming knowledge of social worker’s intentions. |
| Cognitive and knowledge-based trust | Breakdown of trust … due to loss of understanding of partners, wrong expectations, lack of knowledge about tasks and limitations. Trust comes from transparency and knowledge and lead to consent to involving the support worker in the task. Expectation and perception in task. Observation reinforces trust. |
| Negotiation of tasks, boundaries, rules. prior to engagement or employment in the task. Seeking to define match and negotiate first. | Physical, role and capability boundaries defined and adhered to. |
| Context and relationship definition.  | Support worker relationship set within a system involving social networks, supervisors, managers, clients. Support worker gains understanding of the social system. |
| Role drift, stepping over the line. Change in expectations. For example, partially-sighted starts to rely and expect too much from support worker  | Role drift results in relational capability loss since support worker cannot fulfil expectations. Responsibilities and boundaries must be reaffirmed.  |
| Loss of trust resulting in to anxiety, disruption of tasks, rejection of robot’s role, failure to achieve task, increased errors and faults, ethical damage etc.  | Support worker anxiety due to being unable to do the job when the role drifts. Nature of this trust, what happens when it breaks down. Requirement of transparency for trust. Evidence-based trust. |
| Importance of communication and detecting cues, exchanges. | Watchfulness, presence, awareness.Instability and uncertainty leads to increased communication, reduction in autonomy, reining in of tasks. |

Table 1 Analysis

**7. Discussion**

**7.1 Narrative as a basis for socially inspired robotics**

The narrative of Mary and the social worker provides an example of a scenario that can be employed in socially inspired robotics. Observing the social interaction, and particularly the communication within a working relationship can provide the basis for identifying capabilities and functionings required for a sociable robot to interact within society. I will briefly explore this in two dimensions. Firstly the characteristics of the relationship and the capabilities and communication required for successful execution of interactions between humans and sociable robots and secondly the societal prerequisites for successful deployment of sociable robots in a complex and uncertain environment.

**7.2Capabilities and functionings**

Mary brings a physical capability to the relationship with the social worker. By bringing the physical capability of sight to the relationship, additional collaborative capabilities are enabled such as the ability to find phone numbers and to read and write case notes. But Mary’s capabilities do not enable functioning to be achieved without their joining with the capabilities of the social worker, who has distinct capabilities in the practice of social work, understanding the law, and communicating with clients, for example. It is the joining together of their capabilities in collaborative capabilities that results in the functionings of social work.

We can transfer these ideas to the scenario of a sociable robot collaborating with a human in a social situation. The robot will have some capabilities which the human does not have, for example the ability to identify and classify faces and emotions or the ability to navigate physically through a social gathering. Equally, the human will have capabilities which the robot does not have which may include knowledge, social presence and authority. The collaborative capabilities enable functioning in social situations.

How could the robot/human collaboration work? We can draw inspiration from reflecting on the relationship between Mary and the social worker. Firstly, there are distinct roles and boundaries. While Mary finds phone numbers and fills in case notes, it is the social worker who knows whose numbers are needed and what should be written in the case notes. There was a distinct definition and awareness of roles and responsibilities which was informed by an awareness of each other’s capabilities. The understanding of capabilities was established from prior knowledge and daily communication.

Secondly there are prerequisite physical capabilities which underpin social capabilities. The physical capability which Mary provides includes sight and driving ability. These enable collaborative capabilities in social work practice. Similarly a physical capability of a socially-inspired robot design would be a necessary pre-requisite for socially appropriate behaviour. Physical capabilities may include not only high definition vision, but facial recognition and mood interpretation. Furthermore a network of receptors would be required by the robot to enable awareness of positioning, a reasonable capability of movement, but particularly movement relative to the other. It is Mary’s physical capabilities which underpin the social capabilities which are necessary for the functioning the social worker wished to achieve. Socially-inspired robotics will draw on an understanding of the relative physical capabilities of the human and robot and how they contribute to social functioning.

The physical positioning of Mary in relation to the social worker was important. Equally positioning the sociable robot relative to the human is important. Here a certain subtlety is required – social intervention must be appropriate and discreet. The physical capability of bodily movement and awareness supports a social embodiment in which the robot determines its social positioning depending on the social expectation of the human.

Thirdly a communication capability is involved. There must be a capability of both the human and the robot to send and receive messages and correctly interpret them. The collaboration of Mary and the social worker to create the social work capability is driven by a continuous flow of communication, both verbal and non-verbal. It is not only a matter of instructions concerning whose phone number to find, but also social signals, of anxiety, impatience, discretion, concern and so on. Additionally the communication capabilities of both verbal and through the interpretation of body position and, critically, touch, constitute the point where failure occurs: “at some point trust broke down.” The breakdown was a breakdown in communication, creating a feedback loop of increasing anxiety on the part of both Mary and the social worker.

Communication breakdown between the human and the robot may equally destroy trust and increase anxiety. In the case of Mary and the social worker, the breakdown was both catalysed by and amplified by the changing expectations of the social worker. In the socially-inspired design of robots, care must be taken to understand and explain the limits of the robots capability and to be open to changing expectations which will have to be managed. The social workers expectations of Mary grew as the worker increased her reliance on Mary. She expected too much and then when that wasn’t delivered, expected too little, and hence she started to check Mary’s work and undermine her capabilities. Similarly in the case of sociable robots, as familiarity is gained, expectations increase and limitations are forgotten. This may put the robot in an untenable position where it is trying to deliver the impossible which is beyond its capabilities. At this point of failure, safety may be compromised and accidents happen. For example, robots in the world’s first robot hotel did not have sufficicnt capabilities to act as sociable robots (Economist, 2019). Establishing appropriate social expectation, and maintaining them through communication is a key design issue (Breazeal,2002, p231).

So our study of the case study of Mary and the social worker inspires algorithmic requirements for the sociable robot. The narrative suggests a range of social capabilities be considered, including the interpretation of facial signals and other cues and the capability of social anticipation. In developing a sociable robot, socially inspired robotics would involve studying a range of human social interactions to derive robot equivalents.

**7.4 Social Practice and Policy**

While the capability approach can tend to focus on individual capabilities such as I have identified are required for Mary and the social worker and hence for the human and the sociable robot interaction by analogy, capabilities may be considered at the broader policy and social level. Technologies become entangled in society such that laws, regulations, culture and ritual are influenced by them and influence them.

The breakdown and restoration of the relationship between the social worker and Mary provides an interesting basis for socially inspired robotics. It suggests that the success of sociable robot in society will depend on phenomena and actions at a societal level, as much as individual responses to robots. The social worker becomes anxious about her job and her diminishing sight. She compensates psychologically by requiring more of Mary. Mary has a steep learning curve to understand the process and language of social work and the computer systems underpinning the processes. The social worker loses sight of the limitations of Mary’s skills and forgets the complexity of the tasks Mary is undertaking: indeed does not understand the difficulties with the computer system because she has never used it. As the social worker’s anxiety increased, she failed to discriminate between her own capabilities and those of Mary. This is of course common in human relationships: if I can do something, it’s easy, why can’t you? The boundary between me and you has faded. The only way forward is through communication, honest discussion of what the limitations are and a restoration of roles and boundaries. It also requires recognition of responsibilities on the part of the social worker.

The story of the relationship between social worker and Mary acts as analogy in socially inspired robotics at the level of social design issues. For collaboration there must be sufficient sharing knowledge. This requires communication of information about the sociable robot. Such knowledge would need to be absorbed by the human interacting with the robot. This will require training education and transparency. The audience for the sociable robot must be identified and recruited into the education process. This will require user manuals and, I would suggest, expected educational standards for those interacting with the robot. There will need to be agreed definition of the parameters and boundaries of the social environment and agreed understanding of the purpose and tasks in which the sociable robot is engaged. Mary and the social worker operated in a confined and defined environment. A limited office space, a set of expected interactions, expected responses to requests, well delineated role for social worker and support worker. Such well-defined constraints and boundaries were essential for the expression of the collaborative capabilities and the achievement of functionalities. These interactions will be under pinned by law, regulations and codes.

The role of Mary was defined in a job specification, assessed at interview and monitored during a probationary period. Tasks and roles are defined which must align with Mary’s capabilities. The role is very specific, the tasks constrained. The same should apply to a sociable robot. There should be proper specification of roles, tasks and interaction with humans. Not only should there be proper specification of roles, but tests of those roles and approval before the robot enters the field. Mary provided a CV which defined her experience and skills and which could be tested. The probationary period enabled the evaluation of how Mary fitted in with the social worker, her tasks and the social and cultural environment.

A robot CV will be a substantial document which is inspected by regulators for its match to the requirements of a particular social role. The robot CV must not only define the physical characteristic of the robot which would include layout of actuators and sensors, but also the behavioural characteristics. The social interactions it is capable of – i.e. its social capabilities should be defined with evidence of tests passed. The robots social competency should then be formally examined.

The robot CV should be matched by a job specification. The specification will define expected social roles, the nature of the resulting collaborative capabilities and the functionings supported by robot interaction. The job specification will also need to define boundaries – where the robot can physically move, what the constraints are. The evidence provide by both the robot CV and the job specification is the underpinned by a job contract which explains why the robot matches requirements and what the justification is for deploying the robot in a particular situation. It will also define regulatory expectations, in the same way as Mary was subject to data protection legislation and child protection legislation. The sociable robot then enters a probationary period, in which suitability in the field is judged as the capabilities are exercised. This would require formal feedback from the humans with which the robot interacts and the analysis of camera, sensor and actuator data returned from the robot. Through the feedback of sensor data, the robot remains connected with the manufacturer and the manufacturer retains responsibility in the same way as an aircraft engine manufacturer retains connection and responsibility for the engine in service.

The evaluation of the robot CV and probationary period must be carried out in the context of industrial and regulatory standards. A wide range of ethical standards exist which may be taken into account (Winfield, 2019). But underlying these must be declared values which relate the role of the robot to human values, particularly human well-being. Do we value the robot as a replacement for human and an economic instrument for saving money and removing jobs? Is the robot an excuse for avoiding the time consuming interaction which people need? Or is the motivation for robot deployment human flourishing and the compensation or extension of human social capabilities which will lead to functionings which enable human flourishing? Some standards already exist. For example, the BS8611 ethical standards chart possible ethical risks for robots including, for example, the risk of loss of trust (BSI, 2016). The focus of BS8611 is on ethical harm (Stahl, 2018). It includes issues of loss of trust, dehumanisation and self-learning systems exceeding their limit.

Loss of trust and inappropriate trust are opposite ends of a spectrum. At one end, the reliance on the robot is such that it is left to its own devices and no monitoring is undertaken. At the other end of the spectrum, there is no trust in the robot technology and the robot is rendered useless and cannot apply its capabilities to support collaborative capabilities and hence achieve functionings, because, like Mary, every action is checked and the suspicion is that everything is being done wrong.

What is required is moderated trust based both on knowledge of how the sociable robot is operating and what the expectations of the humans interacting with the robot are. The support worker analogy illustrates the critical requirement for trust and what happens when it breaks down. This breakdown was driven by changes in the expectation of the social worker accompanied by her increasing general anxiety. As the social worker grew comfortable with Mary, she expected more. These expectations became unreasonable and the resulting impatience raised anxiety levels in the relationship. Such a rise in anxiety could be envisaged in human interaction with a robot. An ethical recognition of the importance of trust needs to be underpinned by practical social frameworks. Besides ethical standards, there is a need for standards of social interaction. These standards should define expected processes and outcomes which address environment, culture and personality. Social policy and legislation should provide for penalties for manufacturers who build robots which do not meet the standards. However, social policy must work at different hierarchical levels. At international and national levels, legislation and regulation is required which defined responsibilities in the design and deployment of sociable robots. But it is not national regulations or social work law which affect the loss of trust in Mary. Rather it is the local expectation within the local practice.

At the local level, the expectations of the social worker might have been moderated is a service level agreement had been defined. A service level agreement (SLA) defines exactly what service can be expected by the customer (Overby et al, 2017). It will define expected service levels, how these will be measured and duties and responsibilities of provider and customer. In Mary’s case it could have been used to define the activities Mary undertakes, and the activities and responsibilities of the social worker. An SLA recognises that the delivery of a service places responsibilities on both the customer and the provider. Unless the customer acts in a responsible way, the provider cannot deliver the service. As such the SLA would set expectations for both Mary and the social worker. Hence when expectations drift, a discussion about and reference to the SLA would restore the equilibrium.

An SLA for a sociable robot could be viewed within the capabilities approach. Its purpose is to define the functionings supported by the robot. The SLA for a sociable robot would not only have to defined the expected capabilities of both the robot and the interacting humans, but also the emerging collaborative capabilities and the resulting functionings. Metrics would need to address the functionings which result from the exercising of the collaborative capabilities. The SLA becomes a dynamic reference document which is the basis for trust. The SLA is a transparent document which is understandable and communicated with all stakeholders who have an interest in the sociable robot.

The social worker narrative suggests a number of issues which need tacking in the deployment of sociable robot in a society, in particular it raise questions concerning how trust can be established and maintained.

1. **Conclusions**

If robots are to contribute to human flourishing and well-being, they must offer capabilities which connect with human capabilities and enable humans to make free choices about what they do in order to flourish. This flourishing must be set in a social context and involve social interaction because human flourishing requires social connection and communication.

Thus it is not enough to design and deploy robots in society which offer physical capabilities to compensate for human frailty. They must also express social capabilities. Understanding and identifying necessary social capabilities will require that design knowledge is draw from the study of human social interaction. Robot behaviour and social algorithms can then be designed based on this knowledge. I refer to this as socially inspired robotics as a parallel to biologically inspired robotics.

By exploring examples of social behaviour we can suggest requirements for sociable robots. It may be more important that robots behave and interact using human social patterns than that they appear human in their morphology and physical characteristics.

Socially-inspired robots will give rise to a wide variety of artefacts which may lead to social change, to the recognition of needs for policy, for regulation and standards, to transparent communication between users and robots providers and the definition of the service that the robot is providing. Socially-inspired robots steps well beyond any design of social algorithms and concentrates on the context and locating of the robot in the individual, community and societal life.

**References**

Alkire S (2005) Why the capability approach? Journal of Human Development, 6(1), 115-135

Anderson M, Anderson SL (2010) Robot be good: a call for ethical autonomous robots. Scientific American October 2010

Barlas Z (2019) When robots tell you what to do: Sense of agency in human- and robot-guided actions. Consciousness and Cognition, 75. 102819

Bradshaw JM, Hoffman RR, Johnson M, Woods DD (2013) The seven deadly myths of autonomous systems. IEEE Intelligent Systems May/June 2013, 54-61

Breazeal CL (2002) Designing sociable robots MIT Press

BSI (2016) BS 8611 – Robots and robotic devices: Guide to the ethical design and application of robots and robotic systems (No. BS 8611:2016). BSI Standards Publication

Byagowi A, Kopacek, P (2010) Using natural human gait for a humanoid robot. 19th International Workshop on Robotics in Alpe-Adria-Danube Region – RAAD 2010 •June 23–25, 2010, Budapest, Hungary

Chen Y, Wu F, Shuai W, Chen X (2017) Robots serve humans in public places – Kejia robot as a shopping assistant. International Journal of Advanced Robotic Systems, May-June 2017, 1-20

Collins H (2010) Tacit and Explicit Knowledge. University of Chicago Press

Coeckelbergh M (2008) “How i learned to love the robot” capabilities, information technologies and elderly care, in Oosterlaken, I and van den Hoven (eds) The Capability Approach, Technology and Design, Philosophy of Engineering and Technology 5, Springer

Danaher D (2020) Robot Betrayal: A guide to the ethics of robot deception. Ethics and Information Technology doi:10.1007/s10676-019-09520-3

Deng E, Mutlu B, Mataric MJ (2019) Embodiment in socially interactive robots. foundations and trends in robotics: Vol. 7: No. 4, pp 251-356. http://dx.doi.org/10.1561/2300000056

Déniz O, Hernández M, Lorenzo J, CastrillÓn M (2007) An engineering approach to sociable robots, Journal of Experimental & Theoretical Artificial Intelligence, 19:4, 285-306, DOI: [10.1080/09528130701208174](https://doi.org/10.1080/09528130701208174)

Denzin NK (2001) Interpretive Interactionism. Sage Publications

Dexter Industries (2019) Raspberry Pi robot that reads emotions. <https://www.dexterindustries.com/projects/raspberry-pi-robot-that-reads-emotions/> Accessed 3 /8/2019

Economist (2019) Why the world’s first robot hotel was a disaster. <https://www.economist.com/gulliver/2019/03/27/why-the-worlds-first-robot-hotel-was-a-disaster>

Fellous, J-M (2004) From human emotions to robot emotions. American Association for Artificial Intelligence. <https://www.aaai.org/Papers/Symposia/Spring/2004/SS-04-02/SS04-02-008.pdf>

Foster ME (2019) Natural language generation for social robotics: opportunities and challenges. Phil. Trans. R.Soc.B.374 : 20180027. http://dx.doi.org/10.1098/rstb.2018.0027

Fong T, Nourbakhsh I, Dautenhahn,K (2003) A survey of socially interactive robots. Robots and Autonomous Systems 42, 143-166

Garnier S (2011) From ants to robots and back: how robotics can contribute to the study of collective animal behavior. In: Meng Y., Jin Y. (eds) Bio-Inspired Self-Organizing Robotic Systems. Studies in Computational Intelligence, vol 355. Springer, Berlin, Heidelberg

Gleeson B, MacLean K, Haddadi A, Croft E, Alcazar J (2013) Gestures from industry: intuitive human-robot communication from human observation. HRI 2013 - Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, 349-356

Grifantini K (2008) The flight of dragonfly robots. MIT Technology Review June 11 2008

# Hall K (2019) Your kids will be glad a UK government-funded robot will be changing your nappy and not them. The Register: <https://www.theregister.co.uk/2019/10/28/uk_invests_in_care_robots_research/> Accessed 29/12/2019

House of Parliament (2018) Robotics in social care Postnote 591 December 2018 Parliamentary Office of Science and Technology, London

Knight H (2014) How humans respond to robots. Building public policy through good design. Brookings Institute

Kosaka M (2013) Gait control for guide dog robot to walk and climb a step. In: Gaol F. (eds) Recent Progress in Data Engineering and Internet Technology. Lecture Notes in Electrical Engineering, vol 156. Springer, Berlin, Heidelberg

Landowska A. (2019) Uncertainty in emotion recognition. Journal of Information, Communication and Ethics in Society. 17(3) 273-291.

McBride N, Hoffman R (2016) Bridging the ethical gap: from human principles to robot instructions. IEEE Intelligent Systems 31(5) 76-82

McMullan T (2016) How a robot could be grandma’s new carer. The Guardian (2016) <https://www.theguardian.com/technology/2016/nov/06/robot-could-be-grandmas-new-care-assistant>

Muller J, Stachniss C, Arras K, Burgard, W (2008) Socially inspired motion planning for mobile robots in populated environments, Proc. of International Conference on Cognitive Systems, 2008

New Atlas (2019) Riba the friendly robot nurse <https://newatlas.com/riba-robot-nurse/12693/> Accessed 3/8/2019

Nisimura R, Uchida T, Lee A, Saruwatan H, Shikano K, Matsumoto Y (2002) ASKA, Receptionist robot with speech dialogue system. IEEE International Conference on Intelligent Robots and Systems

Oosterlaken I (2009) Design for development: A capability approach. Design Issue 25(4) 91-102

Overby S, Greiner L, Paul LG (2017) What is an SLA? Best practices for service level agreements. CIO Magazine 5th July 2017 <https://www.cio.com/article/2438284/outsourcing-sla-definitions-and-solutions.html> Accessed 2/8/2019

Pavic M (2010) Robo spiders are multi-legged mechanical marvels Wired <https://www.wired.com/2010/04/gallery-spider-robot/>

Ramirez OA, Khambhaita H, Chatila R, Chetouani M, Alami, R. (2016) Robots learning how and where to approach people. 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) August 26-31, 2016. Columbia University, NY, US

Science Focus (2019) Wild ideas to blow your mind December 2019

Sen A (1999) Development as freedom. New York Anchor Books

Stahl B, McBride N, Wakunuma K, Flick C (2014) The empathic care robot: a prototype of responsible research and innovation. Technology Forecasting and Social Change, 84, 74-85

Stahl BC (2018) Implementing responsible research and innovation for care robots through BS 8611. In: Bendel O. (eds) Pflegeroboter. Springer Gabler, Wiesbaden

Suchman L, Trigg R, Blomberg,J (2002) Working artefacts: ethnomethods of the prototype. British Journal of Sociology 53(2) 163-179

Suchman L (2006) Reconfiguring human-robot relations. 15th IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN06), Hatfield, UK. September 6-8 2006

Suchman L (2007) Human-machine reconfigurations: Plans and situated action (2nd Edition) Cambridge University Press

Turkle S (2010) In Good Company. In Y. Wilks (Ed.), *Close engagements with artificial companions*. Amsterdam: John Benjamins Publishing.Willemse

UK Government (2019) Care robots could revolutionise UK care system and provide staff extra support <https://www.gov.uk/government/news/care-robots-could-revolutionise-uk-care-system-and-provide-staff-extra-support> accessed 18/12/2019

Vertesi J (2012) Seeing like a Rover: Visualisation, embodiment and interaction on the Mars Exploration Rover Mission. Social Studies of Science 42(3) 393-414

Winfield AFT, Jirotka M (2018) Ethical governance is essential to building trust in robotics and artificial intelligence. Phil. Trans. R.Soc. A376

Willemse CJAM, Toet A, van Erp JBF (2017). Affective and behavioral responses to robot-initiated social touch: Toward understanding the opportunities and limitations of physical contact in human–robot interaction. *Front. ICT* 4:12. doi: 10.3389/fict.2017.00012

Winfield AFT (2019) Ethical standards in robotics and AI. Nature Electronics 2: 46-48

Ye C, Ma S, Wang Y, Li,B (2003) Coupled-Drive-Based joint dsign of a snake robot and its body lifting method. Proceedings of the 2003 IEEE Changsha, China - October 2003 International Conference on Robotics, Intelligent Systems and Signal Processing