AI & Soc (2016) 31:581–585
DOI 10.1007/s00146-015-0635-2

OPEN FORUM

Investigating and designing social robots from a role-theoretical perspective: Response to "Social interaction with robots—three questions". In Gesa Lindemann (this volume)

Martin Meister[1] · **Ingo Schulz-Schaeffer**[1]

Received: 31 October 2014 / Accepted: 8 October 2015 / Published online: 29 October 2015

Abstract As proposed by the editors, we give statements to two of the questions formulated above, briefly sketch some basics of a role-based approach to socio-technical innovation and give some examples, as to why a role-based approach might be helpful for observing and contributing to the field of Social Robotics.

Keywords Social Robotics · Sociological role theory · Socio-technical innovation · Institutions of robot use

1 Introduction

In our contribution, we address questions one (degree of autonomy of the robot) and question three (institutionalized forms of the use of robots) from an explicitly sociological perspective, and more specifically, from the point of view of sociological role theory.

Response to the question 3: "Robots beyond the laboratory: A necessity to institutionalize forms of use?"

True, on the one hand "the structured artifact addresses users with particular expectations of how to make use of the respective technologies"; and there are numerous empirical descriptions of how artifacts prescribe human actions. But on the other hand, there are not only user-driven technical innovations and uses of technology that were not intended by the designers, but it is well known in STS that every introduction of artifacts in social settings inevitably alters these settings. Starting with Barley (1986), numerous studies in the fields of sociology of technology and organization studies have shown that the introduction of new technology leads to major changes in the set of professional roles and hierarchies in organizations, e.g., the distribution of professional expertise and power relations between patients, nursing staff, doctors, and technical people in a hospital or nursing home. Both these perspectives on the emergence of more or less stable relations between human actors and artifacts led to the concept of socio-technical networks. As is well known, the basic question here is how to conceptualize the interplay between human actors, social context and artifacts without overstretching the generalized principle of symmetry. We propose role theory as a promising starting point for this conceptualization.

Response to the question 1: "The autonomous robot—a too ambitious aim?"

In contrast to conceptions of autonomy in AI and philosophy, for designers of "social" or "sociable" robots the term autonomy or the distinction of degrees of autonomy refers to the most challenging environment for a technical apparatus: the interaction with humans in their natural environment like private households or nursing homes. These environments are especially challenging because there is a tremendous amount of potentially important variables that have to be modeled by the designers and which have to be sensed by the robots in real time, and in an environment that is dynamically changing. For a robot to be able to handle these conditions, the basic problem is not to be able to perform with higher levels of autonomy,

✉ Martin Meister
martin.meister@uni-due.de

Ingo Schulz-Schaeffer
schulz-schaeffer@uni-due.de

[1] Institute for Sociology, University Duisburg-Essen, Lotharstr. 65, 47057 Duisburg, Germany

but to reduce the complexity of the situation in a way that allows for smooth interaction with the humans in the situation. As we know from sociology, human actors are confronted with the same problem (with even greater "information input" and "information processing capacity," e.g., reflexivity). The solution human society provides for this problem of complexity is, according to sociology, the ability of actors to follow roles that in most cases take the form of generalized expectations, which are adequate for the situation at hand. Perceiving interaction partners via typical roles and sending cues that one is acting according to a recognizable role renders it unnecessary to take all the possibly relevant aspects of individual actors into account and makes it possible to choose actions that fit the normal expectations that are attached to that role. We propose that this principle can be applied to the domain of robotics. For a robot, the ability to play out more or less differentiated roles according to situational requirements is located on a higher level of cognitive skill than just behavioral reaction (or "behavioral autonomy"), but does not require substantial properties of organisms like reproduction or properties of humans like "free self-consciousness." It is a great advantage of a role-theoretical approach to socio-technical innovation that such a comparison of substantial properties (and abilities) is not required because the characteristic of positions is that they are independent of their "incumbents" (Dahrendorf 1968 [1958]: 34; see next section).[1]

Before we give some examples, as to why a role-based approach might be helpful for observing and contributing to the field of Social Robotics, we will briefly sketch some basics of a role-based approach to socio-technical innovation.

2 In a nutshell: the role-theoretical perspective on socio-technical networks

Technological innovations do not only consist of technological artifacts but of sets of interrelated heterogeneous components. Each successful innovation requires its heterogeneous components to become mutually adapted to each other, resulting in a sufficiently consistent and coherent behavior by the constellation as a whole. We suggest viewing the heterogeneous components of a technological innovation as positions. The constellations that constitute innovations should be viewed as position fields, that is, as sets of interrelated positions. We use the terms "position" and "position field" the way they are defined in sociological role theory. According to Dahrendorf (1968 [1958]: 34) a social position is a "place in a field of social relations." Positions are constituted and defined by the patterned relations that exist between them and other positions. Positions are, so to speak, the nodes of the network that results from the patterned relations between theses nodes. A set of positions that are interrelated in this way is called a position field. The patterned relations between positions result from patterned expectations and corresponding patterns of behavior which are called roles. In role theory, a role is "the behavior oriented to these patterned expectations of others" (Merton 1957: 110).

In our view, technological innovations consist of fields of positions to be occupied by human actors or technological objects (cf. Schulz-Schaeffer 2014). Humans act as components of technological innovations by assuming particular positions, for instance, the position of operator or user of a particular technological artifact. Since positions are defined by their patterned relations to other positions, they require their incumbents to comply with the role expectations associated with the position. According to sociological role theory, a core characteristic of positions is that they "may in principle be thought of independently of their incumbents" (Dahrendorf 1968 [1958]: 34). Positions "do not cease to exist when they become vacant" (p. 34) and do not depend on the personality or the existence of particular individuals.

With technological artifacts, there is the same difference between individual and position as with human occupants of positions. Our expectations as users of technological artifacts are based on the assumption that the individual objects we are dealing with are occupying particular positions. Our expectations on their properties and their behavior are derived from this assumption. We expect that these objects in their property as occupants of this or that position will behave in a particular way. Thus, in a sense quite similarly to that of role theory in sociology, our expectations are role expectations, while the corresponding behavior is role behavior. The relation between the single object and the position it occupies as a particular kind of technological artifact is in some crucial respects rather similar to the relation between person and position: The positions as car, light bulb, or as whatever kind of technological artifact do not depend on individual characteristics or on the existence of particular single objects, and they also do not cease to exist when they become vacant. The same applies to the roles linked with the positions. They also exist independently from any particular object.

To propose a technological innovation thus is to suggest establishing a new position field, a new set of interrelated positions. It is a new idea about which positions should exist, how they should be related to each other, which role expectations and role behaviors should define these relations, and which entities—technological objects or human

[1] It could be another conceptual advantage of this independence from concrete "incumbents" that individual characteristics are not necessary for modeling. Hence privacy is not a critical issue here at least in principle. But we will not address normative questions in this article.

actors—should occupy these positions. In this sense, it is true, as Callon and Akrich have argued, that every technological innovation is necessarily accompanied by the "delineation of a scenario" (Callon 1986: 26) which envisages these roles and role relations (cf. Akrich 1992b: 208). To the degree that the role relations are inscribed in the technological artifact, the artifact itself "can be described as a scenario replete with a stage, roles, and directions governing the interactions between the actors (human and nonhuman) who are supposed to assume those roles" (Akrich 1992a: 174). This is not to say that engineers are necessarily aware of the scenarios implied in their artifacts. Rather, the underlying assumptions, especially the assumptions about the context of use are often to some extent tacit assumptions. Of course this does not mean that the scenarios which are explicitly envisaged or tacitly presupposed by the engineers are the same scenarios that eventually become reality. The developing innovation may be subject to a lot of different and competing influences, and thus the underlying scenario may change accordingly.

3 Examples from the field of Social Robotics

We now turn to robotics again, to briefly illustrate how these role-theoretical considerations can be applied to the emerging field of Social Robotics—in a way that is not a purely external perspective on robotics, but takes its starting point at discussions in the field itself.

If robots leave the laboratory, the first question is: In which kind of context are they placed? Industrial robots need a carefully designed context to be able to fulfill complicated but always identical tasks. Field robots, the ones with the highest degree of autonomy in technical terms, are sent to settings which have to be as far away as possible from humans: deserts, the deep sea, or Mars. Finally, service or "social" ("sociable" or "socially interactive") robots (see for these definitions: Fong et al. 2003; Dautenhahn 2007) find the limits of their machine autonomy in the needs and the safety of the humans they are built to interact with (for this common differentiation of robotics into subfields, see e.g., the latest version of a leading handbook Siciliano and Khatib 2008).

The emerging field of Social Robotics originated from an association of robotics engineers with an interest in human domains, and scholars from human–computer-interaction research (in which psychological and social sciences approaches have always played an important role) and, in recent years, its subfield human–robot-interaction research (HRI). As mentioned above, the complexity of the social situations in which the robot has to (inter)act is explicitly addressed as one basic problem in Social Robotics. Comparability of different empirical HRI investigations in different settings is addressed as the methodological side of this problem (for a detailed description of this issue see Meister 2014). One way to reduce the complexity of situations is to draw on generalized expectations, and as such, on social roles as an important class of them.

From a quite traditional sociological point of view, the introduction of a more or less advanced robot into any setting can be investigated as a challenge of given sets of positions and their associated role relations. To illustrate this claim, we do not have to speculate about the introduction of robots that, technically, do not exist yet—it of course makes sense to investigate them by using mock-ups—but to look at more simple artifacts that do exist and are already in use. One striking example is the use of robotic vacuum cleaners. Even this rather prosaic technical artifact leads, according to Sung et al. (2010), with respect to the distribution of roles, to quite drastic changes that can be expressed in role-theoretical terminology: The introduction of "robots changed social roles in the house. They induced collaborations among more householders to complete a manual task that used to belong to one person prior to robot use (e.g., Mom for cleaning). This finding suggests that even for a simple utilitarian robot, the ability to act according to social rules can become critical for long-term acceptance" (ibid: 427). But it turns out that in reality not all human users follow the role that was prescribed to the robot by its designers (the "script" of the robot). The robots are also used experimentally, e.g., thrown from the roof of the house or drowned in the basin, simply to see what happens. The position of the robot is purposefully altered by the human users. So from an engineering stance, Sung et al. (2010: 427) come to the conclusion that the role of the robot in the socio-technical network should be prescribed more narrowly by technical means: "Designers can enhance a safety system that can protect the robot from people's random experiments, such as running it at a place with height" (ibid).

More principally, in HRI research typical patterns of HRI are sometimes modeled as social roles. The expectation is that these patterns should be applicable to settings that strongly differ between domains. This makes it unnecessary to take all the possibly relevant aspects of individual actors, "individual" robots, and specific settings into account. One influential example for this kind of approach is a top-down framework with only five principal roles—supervisor, operator, mechanic, bystander, and teammate—as a basis for empirical evaluation of HRI (Scholtz 2003). Within this framework, incumbents of the same interactional positions can be humans or, in most cases, also robots. The author adds that the first and the second roles are exceptions because they are specific for human control of robots. Of course this small number of

roles can be differentiated further if it is necessary to grasp more details of HRI.

An example for a bottom-up, role-based approach in HRI is the identification of "fundamental patterns of human interaction with the physical and social world" (Kahn et al. 2008: 98), which can be understood as episodes of perception of and interaction with technology that often appear. These episodes can be generalized, given the expectation that they will occur again, if particular cues are perceivable.

Patterns like "initial introduction," "in motion together," "recovering from mistakes," or "reciprocal turn-taking in game context" (some of the patterns observed by the authors in robotics experiments with children) define interactional roles for the humans and the robots involved in the episodes. These empirically identified networks of positions (here: positions in interaction sequences) are thought to form an expandable model kit for designing HRI, with the explicit goal of finding more general roles via abstraction from single empirical cases. Other authors have extended this approach to "cue-oriented design patterns" which start from "interaction cues (or social cues) that a robot can perceive and act upon or express in an interaction" (Kuo et al. 2011: 446). Just as in human social life (the identification of appropriate roles; see above), "reading" such cues correctly would "tell" the robot, e.g., whether and when it is expected to take the role of initiator or responder in a given situation, and what is relevant in that situation and (more importantly) what is not relevant. While addressing a rather technical problem (task analysis), the authors work on a cognitive model of the interaction and thus of the technical design of the robot.

4 Conclusions

From these examples, it seems that the use of the concept of social roles in Social Robotics quite smoothly ranges from the empirical description of changes in socio-technical networks to the inspiration of engineering design. On the one side of this spectrum, there is the more traditional focus on the observation of changes in a position field via introduction of novel role-takers or novel positions. On the other side of this spectrum, the "patterns approach" and especially its extension propose a blueprint for the technical design (the "building") of social (or "sociable" or "socially interactive") robots. The possible application of sociological role theory as an inspiration for the elaboration of concepts of Social Robotics seems to apply across this whole range. This shows, or at least illustrates, that the understanding of sociological role theory proposed here is not only a suitable theoretical tool for observing a field, but also a voice heard in the observed field itself.

But there is also a warning against overstretching symmetry in our proposed use of sociological role theory. All the sketched examples and discussions point, at least in principle, to an exchangeability of the nature (human, technical or other) of incumbents of a position. But even when restricted to cue-oriented role-taking, looking at concrete cases seems to reveal position fields where at least some positions—by purpose or not and at the current state of technological development—are restricted to humans. This has nothing to do with "autonomy," but might be a hint that the range of possibilities for enacting positional roles is still unequally distributed between humans and robots—drastically unequally distributed. Moreover, the promising ability of robots to act as role-takers seems so challenging, that no one so far—in Social Robotics as in the Sociology of technology—has thought about robots as situated role-makers (cf. Turner 1962: 22). This points to an issue that is not yet addressed in robotics: the (preferably active) role-making of artifacts, here: robots.

References

Akrich M (1992a) Beyond social construction of technology: the shaping of people and things in the innovation process. In: Dierkes M, Hoffmann U (eds) New technology at the outset. Social forces in the shaping of technological innovations. Campus, Frankfurt am Main, pp 173–190

Akrich M (1992b) The de-scription of technical objects. In: Bijker WE, Law J (eds) Shaping technology; building society. Studies in sociotechnical change. The MIT Press, Cambridge, MA, pp 205–224

Barley SR (1986) Technology as an occasion for structuring: evidence from observations of CT scanners and the Social Order of Radiology Departments. Adm Sci Q 31:78–108

Callon M (1986) The sociology of an actor-network: the case of the electric vehicle. In: Callon M, Law J, Rip A (eds) Mapping the dynamics of science and technology. Sociology of science in the real world. Macmillan Press, London, pp 19–34

Dahrendorf R (1968 [1958]) Homo sociologicus: on the history, significance, and limits of the category of social role. In: Dahrendorf R (ed) Essays in the theory of society. London, Routledge and Kegan Paul, pp 19–87

Dautenhahn K (2007) Socially intelligent robots: dimensions of human–robot interaction. Philos Trans R Soc Biol Sci 362(1480):679–704. http://rstb.royalsocietypublishing.org/content/362/1480/679.full.pdf+html. 22 Nov 2012

Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42: 143–166. http://vrai-group.epfl.ch/papers/RAS03-TF-IN-KD.pdf. 14 Aug 2003

Kahn PH, Freier NG, Kanda T, Ishiguro H, Ruckert JH, Severson RL, Kane SK (2008) Design patterns for sociality in human–robot interaction. In: Proceedings of the 3rd ACM/IEEE international conference on human robot interaction (HRI '08). ACM, New York, NY, pp 97–104. http://delivery.acm.org/10.1145/1350000/1349836/p97-kahn.pdf. 27 Mar 2013

Kuo IH, Jayawardena C, Broadbent E, MacDonald BA (2011) Multidisciplinary design approach for implementation of interactive services: communication initiation and user identification for healthcare service robots. Int J Soc Robot 3:443–456

Meister M (2014) When is a robot really social? An outline of the robot sociologicus. Sti-Studies 10(1), 85–106. http://www.sti-studies.de/ojs/index.php/sti/article/view/145. 24 Jan 2014

Merton RK (1957) The role-set: problems in sociological theory. Br J Sociol 8(2):106–120

Scholtz J (2003) Theory and evaluation of human robot interactions. In: Proceedings Hawaii international conference on system science (36). http://www.hicss.hawaii.edu/HICSS36/HICSSpapers/ETACH01.pdf. 14 Aug 2013

Schulz-Schaeffer I (2014) The position fields of technology: a role-theoretical approach to socio-technical networks

Siciliano B, Khatib O (eds) (2008) Handbook of robotics. Springer, Berlin

Sung JY, Grinter RE, Christensen HI (2010) Domestic robot ecology: an initial framework to unpack long-term acceptance of robots at home. Int J Soc Robot 2(4):417–429

Turner RH (1962) Role-taking: process versus conformity. In: Rose AM (ed) Human behavior and social processes: an interactionist approach. Routledge and Kegan Paul, London, pp 20–40