# Introducing a Methodological Approach to Evaluate HRI from a Genuine Sociological Point of View

Diego Compagna(✉), Manuela Marquardt, and Ivo Boblan

Technische Universität Berlin, FG Regelungssysteme EN11, Einsteinufer 17, 10587 Berlin, Germany
{diego.compagna,ivo.boblan}@tu-berlin.de,
manuelamarquardt@gmx.de

**Abstract.** The evaluation of human-robot interaction (HRI) is still a major methodological challenge. Despite the interdisciplinary nature of the field, sociologically inspired contributions are still rare. This paper aims to introduce a theory-driven method according to a sociological interaction concept to evaluate HRI and identify aspects of successful and satisfying interaction experiences. It combines Harold Garfinkel's breaching experiments with a frame analysis inspired by Erving Goffman. Sociologically, the method relies on a definition of social interaction based on the symbolic interactionism paradigm.

**Keywords:** Breaching experiments · Ethnomethodology · Frame analysis · Symbolic interactionism · Creativity engineering

## 1 Introduction

The normative criteria for a successful interaction in the field of HRI usually relates to the interaction being pleasant for the human being involved – without specifying further how this pleasure is defined or measured. A common goal is to create an interaction experience similar to the interaction with other human beings. Reflecting these trends, two questions are of major interest: Firstly, which factors are important in evaluating an interaction experience (reflecting critically on the assumption that it should be fluid and smooth or as "natural" as possible)? And in direct connection with this, secondly, to what extent and in which situations is an interaction similar to human-human interaction preferred? With regard to this second question, the field of sociology – and especially sociological theory – is promising in identifying factors constitutive of human-human interactions and transferring them situationally to human-robot interactions.

In this paper, we want to present the outline of a theory-driven method for evaluating HRI, which can address both questions at once. To this end, we primarily consulted Erving Goffman's "Frame Analysis" [1, 8], presented in his "Microstudies on Social Interaction" [2, 3] combined with Har-old Garfinkel's "Ethnomethodology" and breaching experiments [4]. Both authors assume that every social interaction is shaped by situation-specific – and therefore context dependent – expectations. Goffman primarily focuses on how these expectations can kept stable over time, whereas Garfinkel is primarily interested in the mechanisms (or methods) that the interacting


J.T.K.V. Koh et al. (Eds.): Cultural Robotics 2015, LNAI 9549, pp. 55–64, 2016.
DOI: 10.1007/978-3-319-42945-8_5

entities use to negotiate an alignment of expectations and to elucidate to one another how their actions should be understood. Garfinkel used breaching experiments to determine these ethno-methods. Combining these with a frame analysis, we are endeavoring to develop a method that is independent of particular cultural contexts, as it operates within the culturally defined boundaries of what is commonly considered to be a functioning interaction. This is especially important in light of the fact that comparative studies for Europe and Japan have reached the conclusion that the concepts of robot agency, as well as an appropriate user-robot interaction, differ significantly [5–7].

To introduce our sociologically inspired methodological HRI evaluation approach, in Sect. 2 we present general assumptions regarding HRI from a sociological point of view. The field of biomimetic robotics is a research area that revolves around human-centered construction and design – which therefore provides a promising application context for our method. The general assumptions of biomimetic robotics are discussed in Sect. 3, whereupon we want to show the potential of sociologically inspired evaluation methods for HRI, comparing two studies that used similar methods and outlining the advantages (as well as probable disadvantages) of our proposed method, combining breaching experiments with frame analysis (Sect. 4). The paper closes with some concluding remarks (Sect. 5).

## 2 General Assumptions Regarding HRI from a Sociological Point of View

Most of the research on HRI adopts a psychological view and therefore selects –in sociological terms – a methodologically individualistic approach (see e.g. [10, 11], and most of the paper presented in [12]). These approaches focus on the individual and disregard socially constructed reality. Genuine sociological contributions – using models, definitions, or theories of social (inter-)action – are still rare (see e.g. [13, 14]).

We use an interactional approach that procedes from George Herbert Mead's concept of symbolically mediated interaction [15]. Interaction is symbolically mediated because the meaning of a symbol (e.g. what is the right gesture for a greeting) is negotiated in an interaction between ego and alter. The meaning of a symbol depends on the reaction of alter on an action of ego and is therefore determined ex post – although intersubjectively between the two subjects ego and alter. Social reality and social meaning are the effect of successful interactions between social actors, as a reaction of alter is always related to the prior action of ego and is what gives ego's action meaning. In the end the understanding I have of myself also derives from the way that I see myself from others' point of view.

Within the scope of sociological work and thought, adopted theories and approaches are connected with the genealogic motivation of criticizing Talcott Parsons' predominant Structural Functionalism from the late 1950s up to the early 1970s. The main goal was – and in part still is – to show how social order is (literally) created by individuals. A systemic view of society and social reality puts the individuals in the shadow of functional necessities, normative expectations, and predefined role sets. Garfinkel's and Goffman's theories opposed this view of society. Their strength relied in proponing theories that could explain the emergence of social order as a bottom-up phenomenon (and not the other way around). However, there are major differences between these two approaches. Garfinkel's assumes an extreme position with regard to the importance of the subject and his or her performance to establish a stabile social system of reference. Goffman on the other hand assumes a position in the middle ground between interactionist theories and system theories, insofar as he relies heavily upon the concept of role, role sets, role expectations and so on.

Garfinkel's approach, ethnomethodology, begins from the very radical assumption that in every interaction, social actors have to readjust their understandings, beliefs, etc. regarding a situation. What actors have in common are the methods that they use to achieve that goal. In contrast, Goffman builds his theory of the relevance of frames on the assumption that a certain amount of meaning is pre-established by prior interactions. Even if the meaning of actions, words, and other symbols relies on the actors' re-affirmation, the main effort for them is to learn the specific codes of the culture they live in and interpret them correctly. The next goal is to determine the frame of reference in every given situation to quickly interpret the actions of alter ego successfully. The two theories are similar in a broader view of social theories. They both consider social reality as the result of the interpretation and affirmation (or establishment) of the meaning of actions with respect to words/symbols. They both rely on a constructivist view of sociality. In this regard, they are easily transferable to HRI settings. However, they represent one cornerstone of the possible assessment of a typical HRI situation. If HRI is seen as an equivalent to Human-Human-Interaction (HHI), these theories could be adopted to analyze the situation in terms of robots being participants in the construction of social reality. The benefit of using these theories would be to address the consequences implied by most social robot developments. To realize a HRI that is very similar to a HHI means inviting robots as equal social partners in the construction of what humans have (so far) claimed to be exclusively their product. On the other hand, using these theories could also reveal to what extent HRI is and should be similar to HHI. Adopting them to understand the details of HRI is the litmus test for the precise amount of genuine sociality involved. In newer concepts sociality isn't limited to humans. The agency of non-humans is taken into account in the Actor-Network-Theory, for example. Nonetheless, using two traditionally constructivist theories that focus purely on humans as social actors is still legitimate because (the development as well as the study of) social robotics is explicitly based on this very grounds.

These assumptions, presented here very briefly, are crucial for the question at stake: In terms of HRI, they become extremely important for knowing who or what could be an appropriate social actor and serve as alter for ego's interactions. Whether reality and identity can emerge from interaction depends on an entity's capability of being alter. This frame is of paramount importance in establishing a clear definition, especially when related to an interaction with a robot. As a human, alter is an entity that is usually seen as a fully social actor who has all the skills needed to recognize an action from ego as an offer. This capability provides a reaction that could be identified as the act of drawing a distinction. In this case, the interaction proceeds smoothly. If alter is a machine, ego will not expect to be able to create social meaning with it. If the robot's reaction is not the one I expected, it will not be able to alter my identity, beliefs, or definitions of reality.

As long as machines lack the ability to participate in symbolically mediated interaction, they won't be attributed the status of fully social actors.

In many cultures today, humans are the only entities who qualify as social actors [16, 17, 18]. This is not surprising in plural, complex modern societies, as it means a reduction of complexity to define entities capable of being social actors merely as humans. Otherwise ego would constantly have to decide if alter is a proper interaction partner and can react in an adequate way in order to build a common social reality, including identity and horizons of indisputable facts and meaningful questions. Gesa Lindemann states that social actors usually don't need a special indication or a referee to know which entity is suitable to interact with, because everyone knows that humans are valid social actors. For this reason, framing is vital in establishing an HRI setting that is suitable for analysis with breaching experiments: If the status is unambiguous and ego is fully aware that alter is just a robot or a machine, ego can deal with failures of communication and interaction flaws. If the status is not completely clear, if ego doesn't know whether the robot is just a machine (as would be clear for example with a washing machine) or if the robot actually possesses skills similar to human skills, the interaction is deeply disturbed. The possibility of implementing a breaching experiment is strictly linked to this observation and therefore to the correct framing of the whole situation. Conducting HRI research in different cultural contexts can benefit from breaching experiments with proper frame analyses, but only if symbols and triggers for crises are implemented cautiously [19].

## 3 General Assumptions Regarding a New Paradigm of Compliant Robots for Human-Centered HRI

The method presented here is highly suitable to be adopted with robots developed in the rather new paradigm of compliant robotics. This is because compliance is a central requirement for human-centered HRI and therefore approaches an HRI similar to HHI. Interesting approaches have come from the field of bionics or biomimetic robotics. The basic motivation behind the transfer of biological solutions to technological applications is the assumption that optimized biological structures have developed over the course of 3.8 billion years of evolution. To date, over 2.5 million different species have been identified and specific characteristics described to a great extent. Thus, in terms of biomimetics there is an enormous pool of ideas available for solutions to technical problems.

Biomimetics is the application of research and development approaches to technological applications that use knowledge gained from the analysis of living systems to find solutions to problems, create new inventions and innovations, and transfer this knowledge to technological systems. The idea of transferring biological principles to technology is the central element of biomimetics. A commonly accepted definition of biomimetics is:

"Biomimetics combine biology and technology with the goal of solving technical problems through the abstraction, transfer, and application of knowledge gained in interdisciplinary cooperation from biological models" [20].

A biomimetic solution usually includes several steps of abstraction and modification with regard to the specific biological solution at hand and is characterized by the creative transfer of ideas. The field is genuinely interdisciplinary, as biologists and engineers, as well as physicists and computer scientists are involved. In biomimetic robotics, biomimetic solutions are applied similarly in the design, control, and operation of robots. Potential applications are lightweight design, flow optimization, or animal-like behavior in navigation, motion control, and decision-making.

A common definition of a biomimetic robot is:

"A robot in which at least one dominant biological principle has been implemented and which is usually developed based on the biomimetic development process" [21].

The benefits gained from the use of biomimetic robots can be derived from inherent physical properties as well as from biomimetic-based "behavior." These two components complement one another. In biologically inspired robotics, some objects of study have proven to be particularly interesting. In these cases, the focus is on biological concepts (such as neural networks) or physical concepts (such as energy storage), depending on the initial perspective. These objects of study, further classified in terms of biological principles, include the following:

- Energy storage and recovery
- Structure and lightweight design
- Efficiency and power-to-weight ratio
- Neurobiomimetic feedback control
- Adaptive behavior /neural networks
- Sensor fusion
- Complex kinematic chains
- Protection /self-protection /protection of others

The mechanics of robotics deals with the motion of bodies (kinematics) and establishes relationships between the motions, mass, and forces acting on a body (dynamics). Current industrial robots are characterized by rigid kinematics that allow precise control using known methods. Since these robotic systems are powerful and fast, they fulfill their tasks very well in industrial environments, but they cannot be used at the present time for human-machine interaction without implementing additional measures. Properties such as strength, quickness, and precision lose their meaning in this case because the humans to which the robots need to adapt (and not vice-versa) are only moderately strong and move relatively slowly and with little precision. Instead, properties such as lightness and compliance (passive, active) become more important and are primary subjects of research at the present time.

The new paradigm is therefore promising in developing robots with whom HRI is not only successful, but satisfying. Following the assumption that HHI is the best reference for HRI one question to what extent the robot should be humanoid or humanized. One assumption is that the more biological principles are implemented in a biomimetic robot, the more the robot approaches its biological role model in its behavior and properties. As fully humanoid robots are not technologically possible at this point in time, an evaluation might be undertaken to evaluate certain human-like aspects. Examples might include a hand shake as a greeting or the passing of objects between a humanoid robot hand and a human hand. Possible influences in this scenario would be optical properties (e.g. five fingers) or haptic properties (e.g. compliance in the movements). One hypothesis that could be tested in HRI scenarios with breaching experiments and frame analysis is that the more similarities the robotic and the human hand have, the more successful and satisfying is the interaction.

## 4 Evaluating the Quality of HRI with Breaching Experiments

Several studies have made use of breaching experiments or similar experimental setups in HRI research [9, 22–29], but without developing a systematic approach with the goal of establishing a generic evaluation method for HRI. There are two main reasons behind breaching experiments' suitability for HRI settings. Firstly, breaching experiments operate on a high level in terms of the demands of social interaction. Secondly, the results aren't biased by social desirability – as is the problem with most survey and interview methods.

In ordinary HRI experiments, the experimentees are asked several questions after the performed interaction about the quality and subjective impressions of the interaction experience. Compared to HHI, HRI is often disappointing, as the experimentee has to carry out most of the interaction sequence. The human fills the gaps created by the robots' inabilities, which in contrast often leads to a more positive assessment of the experienced interaction. This is due to the fact that many experimentees try to emphasize their own efforts and frame their behavior as a response to the researchers' expectations.

With breaching experiments, one indicator for the success of an interaction is the adoption of repair strategies in case of flaws in interaction. These repairs can be observed and will only be adopted if the interaction is seen as worthy of repair.

These assumptions have to be combined with a situational frame: Muhl and Nagai [9] adopted a similar approach in their study and were able to identify six different strategies that experimentees adopted to repair an interaction with a robot. They used a deception strategy, which led to a reframing of the situation. In a laboratory experiment, they invited people to show objects to a robot and teach it some tasks related to these objects. The robot (or agent) was an animation on a screen and appeared as a baby face with eyes, eyelids, eyebrows, and mouth that could move expressively [30]. It had a biologically inspired saliency mechanism and followed the most important features in the scene with its gaze [31]. This is how the robot could display attention or address its human interaction partner without acoustic sensors or speech processing systems [9]. The deception strategy consisted in directing the robot's gaze temporarily in a wrong direction. Even though the interaction is rudimentary, the experimentees' repair strategies showed a certain belief in the robot's interaction capabilities. They tried to re-attract the robot's attention, for example by pointing to the object, showing it to the robot more closely, or making noise [9].

The interaction crisis was induced systematically. Ego applied cognitive framing about the state of its interaction partner alter, which is in turn a selection of how to approach alter. After a crisis arises, ego tries to repair the frame as long as it seems worthwhile. When realizing the breach, the frame is changed and no more attempts are made to repair the interaction.

Awareness of the frame is important in terms of another aspect: As a researcher, one can achieve a high degree of transparency by checking the realizability of the breaching experiment itself. Muhl and Nagai [9] showed that it is possible to achieve meaningful results in a laboratory setting. By comparing these results with research conducted in a stationary care facility for elderly people [14], Compagna & Muhl showed that the instrument can have different effects in different settings. In an everyday situation, people did not put forth any effort to repair an interaction sequence with a service robot. In contrast to this, a test person suffering from dementia went through the interaction sequence perfectly. A service robot had the task of serving a drink to the inhabitants, addressing the person by speaking [32, 33]. Most people didn't reply to the robot at all or preferred to talk to the other people present. Even if they accepted the drink and the robot thanked them, they didn't respond. This was due to the robot's inability to react flexibly, e.g. by turning a rejection into a request. In this situation, the reaction of alter cannot be semantically constructed as following an action of ego. Interaction in sociological terms is not only endangered by its possible failure, but doesn't occur at all. In contrast to these cases, the person with dementia saw the robot as a fully social actor and interaction partner, even if the robot failed to react or if the robot's reaction was not what it was expected to be.

Within the setting (and framing) of an everyday life situation, the adoption of breaching experiments constitutes a stricter test and should therefore only be implemented with care. As in the previously mentioned example, interaction experiments in a laboratory setting may be more fruitful, as the invited participants' motivation is higher and the situation is controlled – suppressing the tendency to address by-standers.

There should be no doubt that framing and context of an interaction are of paramount importance. Nevertheless, further research is needed to learn more about central aspects related to the frame in which the HRI is carried out.

## 5 Summary

The presented method of breaching experiments combined with frame analysis is highly suitable as an instrument in evaluating interactions between a human and a robot from a sociological point of view. A main indicator of the quality of the interaction is whether and how an explicitly induced crisis is repaired in a meaningful way. This approach can also highlight differences between individuals [9, 14] or broader cultural contexts.

Framing should also be reflected upon for the experimental setup. Findings of HRI breaching experiments have to be reanalyzed with respect to framing. If a proper frame analysis is conducted and the frame is set correctly, the findings can be very helpful in determining the quality of an interaction.

When meaningful repair strategies show that the robot is ascribed the status of a social actor, the interaction can be described "as if" it was a social interaction. One peculiarity of HRI is that humans tend to treat robots as if they were entities with the properties of social actors, although they may well be aware of the robots' inability to repair the interaction (one might also say that the robot is not able to process double contingency [16]). The analogy of interacting with a child could be helpful for further understanding the asymmetrical capabilities between the interacting entities. These differences in social interaction between humans and "as if"-social interaction between humans and robots need further consideration in future research. For the present state, we can assume that an "as if"-social interaction is of higher quality than an unsocial interaction.

In conclusion, observable repair strategies are a huge advantage of breaching experiments that can evidence the quality of the HRI, but only if the framing is considered and chosen wisely. Evaluation results can be used to further develop social robots with interactional skills.

**Acknowledgments.** The Research presented in this paper was primarily supported by the German Ministry of Education and Research. In addition, we would like to thank Mollie Hosmer-Dillard for her support and helpful comments.

## References

1. Goffman, E.: Frame Analysis. Harper & Row, New York (1974)
2. Goffman, E.: Interaction Ritual: Essays on Face-to-Face Behavior. Anchor Books, New York (1967)
3. Goffman, E.: Relations in Public: Microstudies of the Public Order. Basic Books, New York (1971)
4. Garfinkel, H.: Studies in Ethnomethodology. Polity Press, Cambridge (1967)
5. Wagner, C.: Robotopia Nipponica: Recherchen zur Akzeptanz von Robotern in Japan. Tectum-Verl, Marburg (2013)
6. MacDorman, K.F., Vasudevan, S.K., Ho, C.-C.: Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. AI Soc. **23**(4), 485–510 (2009)
7. Kaplan, F.: Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. Int. J. Humanoid Rob. **1**(03), 465–480 (2004)
8. Goffman, E.: Behavior in Public Places. Free Press, New York (1963)
9. Muhl, C., Nagai, Y.: Does disturbance discourage people from communicating with a robot? In: 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Jeju, Korea (2007)
10. Wykowska, A., Ryad, C., Mamun Al-Amin, M., Müller, H.J.: Implications of robot actions for human perception. How do we represent actions of the observed robots? Int. J. Soc. Rob. **6**(3), 357–366 (2014)
11. Feil-Seifer, D., Skinner, K., Matarić, M.J.: Benchmarks for evaluating socially assistive robotics. Interact. Stud. **8**(3), 423–439 (2007)
12. Herrmann, G. (ed.): Social Robotics. Lecture Notes in Computer Science, vol. 8239. Springer, Heidelberg (2013)
13. Burghart, C., Haeussling, R.: Evaluation criteria for human robot interaction. In: Proceedings of the Symposium on Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction, pp. 23–31 (2005)
14. Compagna, D., Muhl, C.: Mensch-Roboter-Interaktion – Status der technischen Entität, Kognitive (Des)Orientierung und Emergenzfunktion des Dritten. In: Stubbe, J., Töppel, M. (eds.) Muster und Verläufe der Mensch-Technik-Interaktivität, Band zum gleichnahmigen Workshop am 17./18. Juni 2011 in Berlin, Technical University Technology Studies, Working Papers, TUTS-WP-2-2012, Berlin, 19–34 (2012)

15. Mead, G.H.: Mind, Self, and Society. University of Chicago Press, Chicago (1934)
16. Luhmann, N.: Soziale Systeme: Grundriß einer allgemeinen Theorie. Frankfurt aM, Suhrkamp (1984)
17. Lindemann, G.: Doppelte Kontingenz und reflexive Anthropologie. Zeitschrift für Soziologie **28**(3), 165–181 (1999)
18. Baecker, D.: Who qualifies for communication? A systems perspective on human and other possibly intelligent beings taking part in the next society. Technikfolgenabschätzung - Theorie und Praxis **20**(1), 17–26 (2011)
19. Lutze, M., Brandenburg, S.: Do we need a new internet for elderly people? A cross-cultural investigation. In: Rau, P. (ed.) HCII 2013 and CCD 2013, Part II. LNCS, vol. 8024, pp. 441–450. Springer, Heidelberg (2013)
20. VDI: VDI Guideline 6220, Part 1: Biomimetics – Conception and Strategy. Differences between Biomimetic and Conventional Methods/Products. Verein Deutscher Ingenieure e.V., Düsseldorf (2012)
21. VDI: VDI Guideline 6222, Part 1: Biomimetics – Biomimetic Robots. Verein Deutscher Ingenieure e.V., Düsseldorf (2013)
22. Weiss, A., Bernhaupt, R., Tscheligi, M., Wollherr, D., Kuhnlenz, K., Buss, M.: A methodological variation for acceptance evaluation of human-robot interaction in public places. In: 2008 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2008, pp. 713–718. IEEE (2008)
23. Weiss, A., Igelsböck, J., Tscheligi, M., Bauer, A., Kühnlenz, K., Wollherr, D., Buss, M.: Robots asking for directions: the willingness of passers-by to support robots. In: Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, pp. 23–30. IEEE Press (2010)
24. Sirkin, D., Mok, B., Yang, S., 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, pp. 11–18. ACM (2015)
25. Bauer, A., Klasing, K., Lidoris, G., Mühlbauer, Q., Rohrmüller, F., Sosnowski, S., Xu, T., Kühnlenz, K., Wollherr, D., Buss, M.: The autonomous city explorer: Towards natural human-robot interaction in urban environments. Int. J. Soc. Rob. **1**(2), 127–140 (2009)
26. Alac, M., Movellan, J., Tanaka, F.: When a robot is social: spatial arrangements and multimodal semiotic engagement in the practice of social robotics. Soc. Stud. Sci. **41**(6), 893–926 (2011). 0306312711420565
27. Nagai, Y., Rohlfing, K.J.: Can motionese tell infants and robots 'What to Imitate'? In: Proceedings of the 4th International Symposium on Imitation in Animals and Artifacts, pp. 299–306 (2007)
28. Short, E., Hart, J., Vu, v, Scassellati, B.: No Fair‼: an interaction with a cheating robot. In: 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 219–26 (2010)
29. Takayama, L., Groom, V., Nass, C.: I'm sorry, dave: I'm afraid i won't do that: social aspects of human-agent conflict. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2099–2108 (2009)
30. Ogino, M., Watanabe, A., Asada, M.: Mapping from facial expression to internal state based on intuitive parenting. In: Proceedings of the Sixth International Workshop on Epigenetic Robotics, pp. 182–183 (2006)
31. Nagai, Y., Asada, M., Hosoda, K.: Learning for joint attention helped by functional development. Adv. Rob. **20**(10), 1165–1181 (2006)

32. Compagna, D.: Reconfiguring the user: raising concerns over user-centered innovation. In: Proceedings VIII European Conference on Computing and Philosophy, pp. 332–336 (2010)
33. Compagna, D.: Lost in translation? The dilemma of alignment within participatory technology developments. Poiesis Prax. **9**(1–2), 125–143 (2012)