

Jacqueline Bellon · Friederike Eyssel
Bruno Gransche · Sebastian Nähr-Wagener
Ricarda Wullenkord

Theory and Practice of Sociosensitive and Socioactive Systems

OPEN ACCESS



Springer VS

Theory and Practice of Sociosensitive and Socioactive Systems



Jacqueline Bellon · Friederike Eyssel ·
Bruno Gransche ·
Sebastian Nähr-Wagener ·
Ricarda Wullenkord

Theory and Practice of Sociosensitive and Socioactive Systems

 Springer VS

Jacqueline Bellon
University of Tübingen
Tübingen, Germany

Friederike Eyssel
Bielefeld University
Bielefeld, Germany

Bruno Gransche
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany

Sebastian Nähr-Wagener
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany

Ricarda Wullenkord
Bielefeld University
Bielefeld, Germany



ISBN 978-3-658-36945-3 ISBN 978-3-658-36946-0 (eBook)
<https://doi.org/10.1007/978-3-658-36946-0>

© The Editor(s) (if applicable) and The Author(s) 2022. This book is an open access publication. **Open Access** This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Planung/Lektorat: Cori A. Mackrodt

This Springer VS imprint is published by the registered company Springer Fachmedien
Wiesbaden GmbH part of Springer Nature.

The registered company address is: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Acknowledgements

We would like to thank our project executing agency *VDI/VDE-IT* and especially Constantin von Dewitz for their constructive collaboration, as well as the Federal Ministry of Education and Research for its support, especially Katrin Nostadt and Annette Eickmeyer-Hehn. Cornelius Höller participated in the project as a substitute for parental leave, after previously also contributing as an intern. We would also like to thank the *poliTE* student apprentices and assistants Stefania Reichberg, Viktoria Ehrmann, Sophia Forntheil, Jessica Kräft, Wanda Hartmann, and Nele Zeckel for their productive collaboration. We would like to express special thanks to Barbara Steins and Anke Rabe for their tireless support with administrative matters. Finally, we would like to thank [DIALECTA.com](https://www.dialecta.com) for the translation of the book.

Funding

The project *poliTE – Social Appropriateness for Artificial Assistants* examined the phenomenon of social appropriateness through the lens of ‘cultural techniques’ to understand socially shared contexts of action and their potential transferability to human-machine interactions. *PoliTE* was a joint project by the Center for Cognitive Interaction Technology (CITEC) of Bielefeld University and the Research College “Shaping a humane future” (FoKoS) of the University of Siegen, funded by Department 616 “Interactive Technologies for Health and Quality of Life” (BMBF) under the motto “bringing technology to people” (grant number: 16SV7880K).

Contents

1	Brief Presentation and Key Project Results	1
1.1	poliTE—Social Appropriateness for Artificial Assistants	1
1.2	Summary of research results	2
2	Structure of This Book and Instructions for Use	7
3	What is Social Appropriateness? Voices from 90 Years of Research.	11
4	The FASA model	19
4.1	The «Type of Action, Conduct, Behaviour, or Task» factor	24
4.1.1	Theoretical perspectives on «Type of Action, Conduct, Behaviour, or Task»	25
4.1.1.1	Knowledge organisation and schemata	25
4.1.1.2	Social Information Processing model and Frame Selection model	33
4.1.2	«Type of Action, Conduct, Behaviour, or Task» of socio-sensitive/socioactive artificial assistants	37
4.2	The «Situational Context» factor.	39
4.2.1	Theoretical perspectives on the «Situational Context»	42
4.2.1.1	Defining a situation	42
4.2.1.2	Cultural schemas, cultural intelligence and culture dimension theories for situation understanding	47
4.2.1.2.1	An example of theoretical multitude: More cultural dimensions and typologies.	52

4.2.2	«Situational Context» (culture) and socioresponsive/socioactive artificial assistants	57
4.3	The «Individual Specifics» factor	63
4.3.1	Theoretical perspectives on «Individual Specifics»	65
4.3.1.1	Social cognitive theory	67
4.3.1.2	Social dominance theory	67
4.3.1.3	On the logic of appropriateness as a Decision-Making theory	68
4.3.2	«Individual Specifics» and socioresponsive/socioactive artificial assistants	71
4.4	The «Relations between Interacting Agents» factor	72
4.4.1	Theoretical perspectives on «Relations between Interacting Agents»	73
4.4.1.1	Respect and recognition	75
4.4.1.2	Goffman's frame analysis	76
4.4.1.3	Social relations model	77
4.4.1.4	Expectancy violations model	78
4.4.1.5	Status and power in interactions	78
4.4.1.6	Face-Saving view	78
4.4.1.7	Etiquette engine	80
4.4.1.8	The politeness principle	80
4.4.2	«Relations between Interacting Agents» and socioresponsive/socioactive artificial assistant	82
4.5	The «Standards of Customary Practice» factor	82
4.5.1	Theoretical perspectives on «Standards of Customary Practice»	85
4.5.1.1	Approaches to «social norms» based on theories of game rules	86
4.5.1.2	Rational choice theory and game theory	87
4.5.1.3	Norm activation model	91
4.5.1.4	Manners	92
4.5.2	«Standards of Customary Practice» and socioresponsive/socioactive artificial assistants	96
5	Social Appropriateness Tree Structures	99
6	Application of the FASA Model to Technical Systems	107
6.1	Application example 1 – A robot barista	108
6.2	Application example 2 – Baby schema	109

6.3	Application example 3 – That robot touch	110
6.4	Application example 4 – If a robot comes down the hallway	111
6.5	Application example 5 – Wait for it... Hello!	112
6.6	Application example 6 – CLIPPY	112
6.7	Fictitious application example/thought experiment	114
7	Reflections on Sociosensitive and Socioactive Technology	121
7.1	Social appropriateness and tact	121
7.2	Why social appropriateness in human–machine interactions? . . .	123
7.3	Sociosensitive/socioactive systems as seemingly human?	124
7.4	Advantages of sociosensitive/socioactive technology/ Why and why not?	127
7.5	Technology does not interpret and does not understand	129
7.6	Politeness as blameless deception – <i>fake it until you make it</i>	129
7.7	Sociosensitive and socioactive technology as an enabling condition and cultivation factor of the human ability to judge social appropriateness?	132
7.8	Differences between humans and technology	133
7.9	Orientation according to rules is not the same as orientation among alternative rules	136
7.10	Challenges of technical implementation	137
	Glossary	141
	References	149

List of Figures

Fig. 3.1	Social appropriateness as a spectrum (Watts 2005, xliii)	16
Fig. 4.1	Factors of social appropriateness.	20
Fig. 4.2	Social information processing model. (Adapted from Döpfner 1989, p. 4).	34
Fig. 4.3	The types of selection considered in the frame selection model (Kroneberg 2010, p. 133)	36
Fig. 4.4	Social importance model applied to general BDI architecture (Mascarenhas et al. 2016, p. 939)	60
Fig. 4.5	Schematic representation of decision factors (Weber et al. 2004, p. 284)	69
Fig. 4.6	Classification of literature analysis on social dilemmas. (Modified after Weber et al. 2004, p. 287)	70
Fig. 4.7	The six maxims of the politeness principle. (Figure modified after Leech 1983, p. 150).	81
Fig. 4.8	The norm activation model. (Modified after Schwartz 1977 and Howard 1981)	92
Fig. 5.1	«Type of Action, Conduct, Behaviour, or Task»	101
Fig. 5.2	«Situational Context»	102
Fig. 5.3	«Individual Specifics»	103
Fig. 5.4	«Relations between Interacting Agents»	103
Fig. 5.5	«Standards of Customary Practice»	105
Fig. 6.1	Software politeness rules (Whitworth 2005, p. 359.)	113

List of Tables

Table 4.1	Type of Action, Conduct, Behaviour, or Task	26
Table 4.2	Situational Context.	40
Table 4.3	Individual Specifics	66
Table 4.4	Relations between Interacting Agents.	74
Table 4.5	Standards of Customary Practice	84



Brief Presentation and Key Project Results

1

1.1 poliTE—Social Appropriateness for Artificial Assistants

poliTE—Social Appropriateness for Artificial Assistants examined the phenomenon of social appropriateness through the lens of ‘cultural techniques’ to understand socially shared contexts of action and their potential transferability to human–machine interactions. *poliTE* was a joint project by *CITEC* (Center for Cognitive Interaction Technology) at Bielefeld University and *FoKoS* (Institute for Advanced Studies “Shaping a humane future”) at the University of Siegen, funded by the Federal Ministry of Education and Research, Department 616 “Interactive Technologies for Health and Quality of Life” under the motto “Bringing technology to the people” (grant number 16SV7880K).

The project’s starting point was the prediction that **interactive systems capable of learning** will continue to gain importance in more and more areas of life. For example, the professional and private everyday lives of many people will be affected by the use of digital voice assistants, social robots, and virtual agents. For **interpersonal interactions**, every culture has developed a set of more or less complex rules and techniques governing behaviour, interactions, and **(social) appropriateness** that for example determine when apologies, greetings, congratulations, reprimands, or other social practices and rituals are appropriate, as well as how these interactions should unfold. Such **cultural techniques of behaviour**, and the customs, behavioural norms, and conventions that regulate them are passed on—often in unwritten form—and acquired by individuals over the course of their socialisation. Thus, modes of behaviour are judged as more or less socially appropriate depending on the person, role, situation, context, culture, time, etc. Especially when our everyday lives are increasingly permeated by

assistance technology, we face questions about the relationship between cultural techniques of behaviour and ‘intelligent’ technological systems: **can and should technical systems be equipped with the ability to interact in a socially appropriate manner?**—this was the key focus of the *poliTE* research project.

As part of *poliTE*, an interdisciplinary literature search in natural sciences and humanities was performed to investigate philosophical, empirical, and technological aspects of social appropriateness. The results were then systematized. To gain an overview of the interdisciplinary research and theory landscape, a comprehensive *topic and actor analysis* (TAA) was performed on the genesis, change, and the conditions of possibility and perception of different forms of socially appropriate behaviour and appropriateness judgements associated with them, allowing a set of recurring topics, research interests, and people to be identified. The topic and actor analysis included extensive term searches in literature databases,¹ journal analyses,² and conference analyses,³ as well as a survey (mixed sample of specialized and non-specialized audience) and in-depth free research on selected topics. In this process we accumulated a total of around 5,000 scientific publications dealing with aspects of social appropriateness or judgement thereof, which are publicly available and browsable in the form of a database at <http://www.polite-data.netzweber.de>. In the second project phase, these results were sorted and systematized, and theories from scientific branches such as psychology, philosophy, sociology, linguistics, anthropology, and theories of culture and media were examined and adapted accordingly to identify any potential of epistemologically fruitful transferability to human–machine interactions in general and to technical systems design in particular.

1.2 Summary of research results

A central result of the analysis of these interdisciplinary approaches to the phenomenon of social appropriateness is the FASA model, a model of the **factors (FA) of social (S) appropriateness (A)**, cf. Chap. 4. The FASA MODEL was established with a special focus on potential transferability to human–machine

¹PsycINFO, PSYINDEX, IEEEexplore, Science Direct, Web of Science.

²Techné, Philosophy and Technology, poiesis & praxis, Philosophy of the Social Sciences, Philosophical Psychology, Ethics & Behaviour, Journal of Consciousness Studies, Journal of Politeness Research.

³IEEE proceedings, SPT, EASST.

interactions and to technical systems design. It consists of the following five factors:

- a) **«Type of Action, Conduct, Behaviour, or Task» (Sect. 4.1):** Actions or behaviour (or sequences thereof) are always viewed as socially appropriate as part of (as a concrete realization of) a certain type of action, conduct, behaviour, or task. The terms ‘behaviour’ and ‘action’ are typically used as synonyms in this book. A more precise distinction can be made by stipulating that behaviour is not deliberate but represents an organic or motor response to stimuli, while action is more cognitively targeted.
- b) **«Situational Context» (Sect. 4.2):** Actions and behaviours are embedded within situations—whether an action/behaviour is socially appropriate depends on the situation in which it unfolds, among other things. «Situational Context» refers in particular to the specific time and place of a concrete interaction, which has a specific spatial structure, and can for example be shaped by the degree of formality (intimate, familial, private, semi-private, public, etc.) of the interaction, the typical roles, behavioural/action and status requirements placed on the interacting parties, their typical hierarchies of preference, the participants’ definition of the situation, and so on.
- c) **«Individual Specifics»⁴ (Sect. 4.3):** «Individual Specifics» addresses individual influences on the construction and/or perception of social appropriateness. Among other things, what is considered appropriate depends strongly on the individual characteristics of the interacting parties, such as their physical, psychological, and cognitive constitution and situational state, or their age, gender, etc.
- d) **«Relations between Interacting Agents»⁵ (Sect. 4.4):** In an interaction, the interacting agents do not meet as solitary agents but as actors who are engaged in relating to one another. Social relations between the participants of an interaction can, for example, take the form of cultural or societal relations such as social closeness/distance, status, respect, degree of familiarity, power constellations, etc.

⁴In the terminology of differential psychology, this factor encompasses inter- and intraindividual differences (see Stemmler et al. 2016).

⁵Here, the interacting parties are not necessarily limited to being human actors.

- e) «**Standards of Customary Practice**» (Sect. 4.5): In social practice, there are implicit norms of action and behaviour. In many cases, there are standards of customary practice about what is considered socially appropriate. These can be understood as a set of customs on a spectrum ranging from ‘**specific**’ (possibly divergent) **group customs** (e.g., customs within different families, sports teams, or companies) on the one hand, which in extreme cases may be negotiated once and for all or alternatively may be constantly situationally renegotiated, to ethically justifiable, and hence ‘**general**’, **regulative norms** on the other hand.

Within these factors, there may be sub-factors, which are described as **factor criteria** in our terminology. As discussed above, a situation is for example constituted by the time of the interaction, its environment, the participants’ roles, etc.—in this sense, the factor criteria are what ‘make’ something a situation. The factors and factor criteria themselves are abstract—factors and factor criteria may not be sensually perceptible. By contrast, so-called ‘**observables**’ can be perceived and observed with senses or sensors: a person’s concrete clothing, smell, hairstyle, etc. are sensually perceptible characteristics from which people and potentially also technical systems can infer parameters of factor criteria and factors.

Observables require interpretation. For instance, what does a display reading ‘19 °C’ actually mean in a specific situation? On a thermostat display, most people would consider this a comfortable room temperature, but on a kettle, it would indicate that the water is still far from boiling (and you need to wait a little longer for your tea, for example). The same applies to observables that function as symbols of meaning in the context of social appropriateness. When observables are interpreted by an observer and provide meaningful clues about particular facts or contexts to this observer, we call them ‘**indicators**’.

The five factors of social appropriateness as presented above, together with the many factor criteria and observables/indicators associated with them, determine what is considered socially appropriate in the presence of **complex conditions and interdependencies**. They also influence the construction and performance of socially appropriate behaviour, as well as judgements of appropriateness in **interpersonal** interactions. Regarding **human–machine interactions** and the design of **technical systems**, some of the conditions that apply to the construction and evaluation of socially appropriate types of action, conduct, behaviour, or task are the same, while some others are very different. The **FASA MODEL** allows the

aspects of a human–machine interaction to be inspected to determine which factors are being considered with what realisations, which factors could or should be considered, and which factors cannot or should not be considered.

These interdependencies between the factors of social appropriateness, their criteria, and any observables and indicators can also be structured into other heuristic approaches to the phenomenon of social appropriateness: for instance, **social appropriateness tree structures** are presented in Chap. 5. These tree structures classify the factors, their criteria, and some associated observables/indicators, offering another way to approach social appropriateness besides the **FASA MODEL** and the explanations of individual factors of the model in Chap. 4. Finally, the complex interdependencies mean that the explanations about theoretical approaches to specific factors in this book are closely linked. We refer to other potentially applicable factors and factor criteria when describing the respective theories.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Structure of This Book and Instructions for Use

2

We would like to make a few remarks about the scope and usage of this book:

This Work Does Not Intend to

- list all observable features of certain recognizable social relationships in the context of social appropriateness (in the terminology introduced in Sect. 1.2: we do not intend to list all observables of specific factor criteria of social appropriateness),
- explain *why or how* specifically observable features (observables) indicate certain social relationships in the context of social appropriateness (factor criteria of social appropriateness),
- indicate which concrete behaviours are socially appropriate in specific interaction situations or which properties must concretely be implemented on a technical level to enable a technical system currently being designed to be capable of interacting in a socially appropriate manner.

Therefore, despite some references to empirical research, this work does not, for example, specify what particular speed a robot should choose when approaching a human or what specific greeting sequence would be appropriate in a formal greeting situation. Regarding social appropriateness, specific action indications like these cannot simply be derived from general rules in an interpersonal context, and a universally applicable ruleset for interactions with artificial assistants does not make sense in the context of human–machine interactions.

This Work Does Intend to

- based on the FASA model, provide **orientation guidelines** about which features (observables) could indicate certain social relationships in the context of social appropriateness (factor criteria of social appropriateness) and what factors and factor criteria of social appropriateness should be considered in concrete interaction situations.
- initiate **reflection** on socially appropriate behaviour. In other words, the book at hand intends to exemplify which observables (could) indicate certain social relationships in the context of social appropriateness (factor criteria of social appropriateness) and what behaviour within the framework of the factor criteria and in idealized contexts could potentially be regarded as socially appropriate or inappropriate, without making any claims of exhaustiveness or exclusivity.

This allows the **FASA model** to be used as a heuristic checklist for many application contexts and interaction situations to see which social relations are relevant in the context of social appropriateness (i.e., the individual factor criteria can be used to check whether a factor or a criterion tends to be relevant or negligible in a specific intended interaction situation or general application context). Additionally, it gives a first look at what features (observables) artificial assistants designed to detect and possibly execute socially appropriate behaviour in specific application contexts or interaction situations—in other words, systems designed to be **sociosensitive or socioactive**—might need to detect and process on a technical level. The application examples discussed in this book are intended to serve as prompts for interpretation and should explicitly not be misunderstood as a blueprint for questions of social appropriateness—even when discussing specific application contexts or interaction situations.

In which contexts, situations, instances of human–computer interaction, etc. it would **be** beneficial to incorporate the dimension of social appropriateness to guide the designs of systems and interactions, and under which circumstances it might be better to choose to avoid designing technology sociosensitively let alone socioactively are questions that are not answered here, while the urgency to ask them is strongly emphasized. This book also lays foundations to formulate answers to these questions to a fitting level of complexity; it raises key questions about the design, regulation, and use of systems that intervene in social spaces and prepares the way for answers. With regard to the example of an approaching robot's speed or the type of greeting chosen, while this work cannot provide

exact or even vaguely normative suggestions on how to make concrete decisions, this work can give hints as to when, why, and where the speed of an approaching robot or the performance of a greeting could have an impact on the perceived social appropriateness of an agent's behaviour.

Another approach to the phenomenon of social appropriateness is given by **social appropriateness tree structures** (Chap. 5). In these structures, the observables/indicators listed as examples in Chap. 4 can be tentatively assigned to the factor criteria of specific factors. This gives us graphs that offer **quick access to social appropriateness**: for example, given a few observables/indicators—technically processed by a sociosensitive/socioactive system—the graphs can be used to check for which dimensions of social appropriateness these observables/indicators might be relevant. Conversely, given parameters of factors and factor criteria of social appropriateness, the graphs can be used to check which observables/indicators are relevant to social appropriateness and might therefore need to be processed by a sociosensitive/socioactive system in each case. The tree structures portray graphically how both of the <time> and <intention> factor criteria are associated with two different factors—<time> is associated with the two factors «Situational Context» and «Type of Action, Conduct, Behaviour, or Task» and <intention> is associated with the two factors «Type of Action, Conduct, Behaviour, or Task» and «Relations between Interacting Agents». Thus, the graphs illustrate the **interconnections between factor criteria and factors, and therefore ultimately between the factors themselves**.

About the Structure of the Content of This Book

- In **Chap. 3**, 'What is social appropriateness? Voices from 90 years of research', the phenomenon of social appropriateness is illuminated with a selection of perspectives to establish a preliminary understanding of what is actually meant by 'social appropriateness' and what forms its investigation can take.
- In **Chap. 4** the **FASA model** is explicated using its factors and references to theoretical approaches in the literature.
- **Chap. 5** presents **social appropriateness tree structures** as an alternative heuristic approach: here, the factors, their criteria, and any prominently associated observables/indicators of social appropriateness presented in the model explication are graphically arranged and interrelated.
- In **Chap. 6** to further illustrate the **FASA model** in the context of technical systems, the model is applied to **examples of technical systems** using various case studies and a thought experiment.

- In **Chap. 7** reflections about the potential and the limitations of the model are presented. Without claiming to be exhaustive, these reflections raise a few critical and ethical considerations about sociosensitive and socioactive technology. They can also be understood as self-reflection on our own research activities.
- **Book Backmatter** consists of a glossary of terminology and a bibliography that merely represents an excerpt of the three-year literature search. A comprehensive database of the research results is available at www.polite-data.net-zweber.de.

Finally, to improve the visual readability of the book, some illustrations from the literature have been modified, as indicated in the corresponding captions.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





What is Social Appropriateness? Voices from 90 Years of Research

3

The very first question when considering social appropriateness as a phenomenon is:

How should Social Appropriateness be understood?

How can we define it? What does it encompass? What do we need to consider if we wish to achieve socially appropriate behaviour (whether by humans or non-human entities)? The literature on this topic and philosophical analysis of the concept of social appropriateness have shown that not only is this definition non-trivial, but there cannot be a universal definition at all. To offer a preliminary understanding and an initial conceptual approach, the next section therefore gives a loose collection of investigations of and explanations about the concept of social appropriateness—presented as **a kaleidoscopic overview of social appropriateness**—compiled during our literature search into various theories and concepts, so that we can begin the book with a definition of the field of social appropriateness. The complete list of literature we've researched is publicly accessible as an interactive database at www.polite-data.netzweber.de. In addition, please note: All the following quotations are our own translations.

A Kaleidoscopic Overview of Social Appropriateness

- Social appropriateness for artificial assistants can be evaluated in service providing contexts. (Hegenbart 2019)
- Social appropriateness is made up of the quantifiable average and particular judgements of a perceiver, a perceived, and the relationship between them. (Kenny and La Voie 1984)
- Social appropriateness depends on the framing attributed to the situation by the interacting agents. (derived from frame analysis, cf. Goffman 1974)

- Social appropriateness consists of cooperation and coordination of social behaviour and can be explained in terms of game theory. (derived from game theory, c.f. e.g., Voss 2001; Diekmann 2016)
- Social appropriateness is a component of human behaviour that is oriented according to environmental affordances and can be influenced by a constructed environment. (derived from affordance theory, cf. Gibson 1979/2014, 1982)
- Like any human behaviour, social appropriateness is made up of the valence of the situation and the individual reaction possibilities of the individual. (derived from the discussion of valence, cf. Koffka 1962)
- Social appropriateness is a component of human behaviour generated by the environmental powers acting on it and the specific needs of the perceiver. (derived from the discussion of stimulating character, cf. Lewin 1926)
- Social appropriateness is status-dependent, and status manifests partly through clothing conventions. (derived from Bohn 2000; Esposito 2004)
- Social appropriateness encompasses different functions of politeness, friendliness, and tact that preserve society and interactions. (Luhmann 1996, 2008, 2017)
- Manners primarily serve the social integration of as many people as possible and facilitate the expression of sovereignty. (Goffman 1982a)
- Certain interpersonal manners—forms of propriety—serve the moral education of man. (Kant 1983)
- Through social appropriateness, there is a dependency on the situation within which the action is unfolding, its specific rules of actions, and the identity of the actors. (Weber et al. 2004; following March 1994)
- Socially appropriate behaviour is based on quantifiable cultural intelligence. (derived from the discussion on cultural intelligence, cf. Ang and van Dyne 2008)
- Social appropriateness serves primarily to avoid social dilemmas and make social interactions successful. (derived from game-theoretical approaches, cf. Voss 2001)
- Social appropriateness depends on constitutive rules for normative standards of appropriateness for the realization of social practices. (derived from Searle 2011)

- Social appropriateness depends on normative standards of appropriateness that are constantly renewed by situational social practice. (Poljanšek 2019)
- Social relations and therefore the framing of the rules of social appropriateness are primarily negotiated through the factors of ‘power’ and ‘status’. (status power theory, cf. Kemper 2011)
- Social (in)appropriateness results from maximization of utility/minimization of cost. (derived from rational choice theory, cf. Coleman and Fararo 1992)
- Social appropriateness is or results from the activation of the norms applicable to a situation. (derived from the norm activation model, cf. Schwartz 1977; Schwartz and Howard 1981)
- Social appropriateness/Socially appropriate robot behaviour results from the correct identification of the ability needs of the robots being deployed. (derived from Dautenheim 2007)
- The phenomenon area of social appropriateness contains expectations, aspects of validity, reference to rules, references to role, personal attitudes, and an experience bank. (Busse 2021)
- Social appropriateness arises from the perception and correct interpretation of social cues. (Vinciarelli et al. 2009)
- Social appropriateness results from the correct selection of a frame or script. (derived from models of frame selection (cf. Kroneberg 2010) and social information processing (e.g., Döpfner 1989))
- Social appropriateness serves to ensure smooth communication and manifests in the observance of the politeness principle. (cf. Geoffrey Leech 1983)
- Social appropriateness amounts to saving face, namely the public self-image of the interacting agents. (derived from the face-saving view, cf. Brown and Levinson 1978, 1987)
- Social appropriateness is behaviour consistent with expectations. (derived from the expectancy violations model, Burgoon et al. 1988)
- Social appropriateness manifests as pragmatic competence in following politeness rule. (derived from the rules of pragmatic competence, cf. Lakoff 1973)
- Manners reinforce (and partly arise from) class affiliations and are part of the habitus. (Bourdieu 1982)

- Social appropriateness encompasses culturally sensitive behaviour. (Höbller 2019)
- Knowledge of social appropriateness contain: a description of certain types of behaviour in the context of social interaction; a stipulation of whether these types of behaviour should be reproduced or avoided; a stipulation and description of specific types of situation in social interactions where the behaviour should be considered desirable or undesirable; a stipulation of the social group(s) in which the rule(s) should apply; a stipulation of situation-related role(s) (rank, status) to which a particular rule of socially appropriate behaviour applies; a specification of the dependency relationships between the various subaspects (from the report by Busse 2019)
- Social appropriate behaviour for technical systems can be constructed through a 5-step design process that in particular reflects the area of application of the system and the users' expectations. (derived from design for etiquette-sensitive HRI application, cf. Bickmore 2001; Miller 2004; Zhang et al. 2001)
- Social appropriate behaviour of technical systems can be constructed with a politeness algorithm based on the face-saving view of Brown and Levinson. (derived from the ideas of the etiquette engine, cf. Wu et al. 2011)
- Social appropriateness encompasses a spectrum ranging from rude to over-polite behaviour, is negotiated in interactions, and depends on the perception of the interacting agent and their expectations of appropriateness mediated by their habitus and the social field of the interaction situation. (derived from Locher and Watts 2005)
- Society defines what is considered expectable expectations in various contexts; every communication and every performance of socially (in) appropriate behaviour contributes to the realization of society. (freely adapted from Niklas Luhmann)
- Socially appropriate behaviour depends on environmental influences, cognitive, emotional, and biological processes, and is mediated by social learning. (derived from social cognitive theory, cf. Bandura 1989)
- Cultural schemas can contain scripts or procedure schemas that specify the socially appropriate sequence of actions. (Schank and Abelson 1997)

- Social appropriateness depends on how much contextualization is associated with individual interactions in a society. (derived from the discussion of low-context and high-context cultures, cf. Gudykunst 1983)
- Event associated with specific rules of social appropriateness are stored schematically in human memory (Bartlett 1932). Schemas are organized by assimilation and accommodation. (Piaget 1936, 1976)
- Judgements of appropriateness depend on the personal preferences of the judging parties regarding group dynamics. (derived, e.g., from social dominance theory, cf. Pratto et al. 1994; Pratto and Sidanius 2012; Ho et al. 2012; Pratto et al. 2016)
- The status of a person influences their behaviour options with regard to the judgement of the social appropriateness of behaviour (e.g., Cornelia Bohn 2000; Elena Esposito 2004), as well as their own judgements of the appropriateness of the modes of behaviour of others.
- In premodern societies, social appropriateness was based on situational and cross-functional status difference (e.g., Bernhard Giesen 1991; David Diehl and Daniel McFarland 2010; Randall Collins 2000). It was recognized, e.g., by clothing regulation, forms of address, and spatial distance and was mediated and fixed by rituals and ceremonies. (Jörg Berns and Thomas Rahn 1995)
- In certain social contexts, such as care, gastronomy, or diplomacy, describable standards of social appropriateness are required. (freely adapted from Youssef 2019)

Finally, the phenomenon of social appropriateness can be presented as a spectrum, as illustrated in Fig. 3.1 in the context of linguistic politeness research (Watts 2005, xliii), a topic that we will discuss in more detail later. This spectrum ranges from grossly impolite (*rude*) to excessively polite (*over-polite*) behaviour and can be divided into three sectors: rude, impolite, and over-polite behaviours are typically classified as inappropriate and are therefore described by Watts as *negatively marked behaviour*. In the region of merely non-polite behaviour, behaviours are typically not noticed for reasons of social appropriateness (*unmarked behaviour*)—in a certain sense, this captures appropriate behaviour in the narrow sense as behaviour that ‘just passes’ (cf. Poljanšek 2022, own

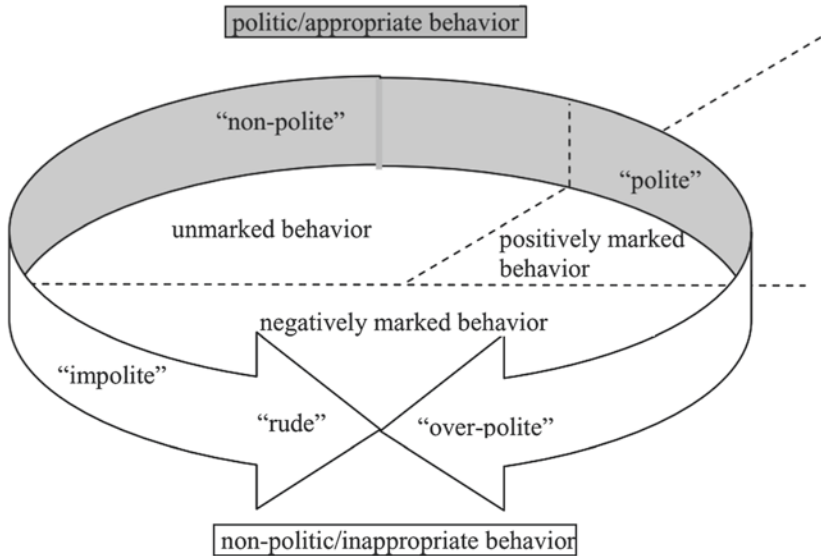


Fig. 3.1 Social appropriateness as a spectrum (Watts 2005, xliii)

translation). Finally, the third sector encompasses courteous (*polite*) behaviour that typically stands out to participants as positive (*positively marked behaviour*).

For these and other approaches to clarifying the question of what structures socially appropriate behaviour and judgements of the appropriateness of behaviours, as well as countless individual aspects and references to technical systems, see also the database of literature on social appropriateness at www.polite-data.netzweber.de.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





The FASA model

4

We don't all carry an etiquette handbook, but everyone seems to know good manners. Although violations exist, most people are polite most of the time. The level of conformity is striking, a fact blurred by vivid memory for occasional lapses. Politeness is ubiquitous, and it's practised automatically. Communities encourage it, and the rules are a centrepiece of childhood socialization. Politeness, even more than early mastery of letters and numbers, is a genuine mark of an educated child, as any kindergarten teacher can attest. (Reeves and Nass 1996, p. 19)

Surprisingly, human beings usually have the ability to 'read the room' and strike the 'right tone' in interaction situations and—without ever referring to a rulebook—know what is considered (in)appropriate in specific situations. And even though the understanding of what is and is not socially appropriate in a certain situation varies among interaction participants, people have an outstanding ability to coordinate with one another very quickly and effectively repair any 'breaks' in communication that may have occurred.¹ As a key result of *poliTE*, we propose five factors that can be considered to analyse the phenomenon of social appropriateness: **a) «Type of Action, Conduct, Behaviour, or Task», b) «Situational Context», c) «Individual Specifics», d) «Relations between Interacting Agents»,² e) «Standards of Customary Practice».** In the presence of complex interdependencies—**represented by the circular**

¹These processes are subject of, for example, psychological, sociological, philosophical and linguistic research among fields such as the cognitive and neuro sciences, communication theory and many others, see for example for notions such as *Repair Strategies/Self-repair* Bolden (2011); Di Venanzio (2016); Sacks et al. (1977); for an approach from a point of view of action and systems theory Parsons et al. (1951); Parsons and Turner (2005); for the notion of *Intentional Attunement* Gallese (2006).

²Here, the interacting agents are not necessarily just human actors.

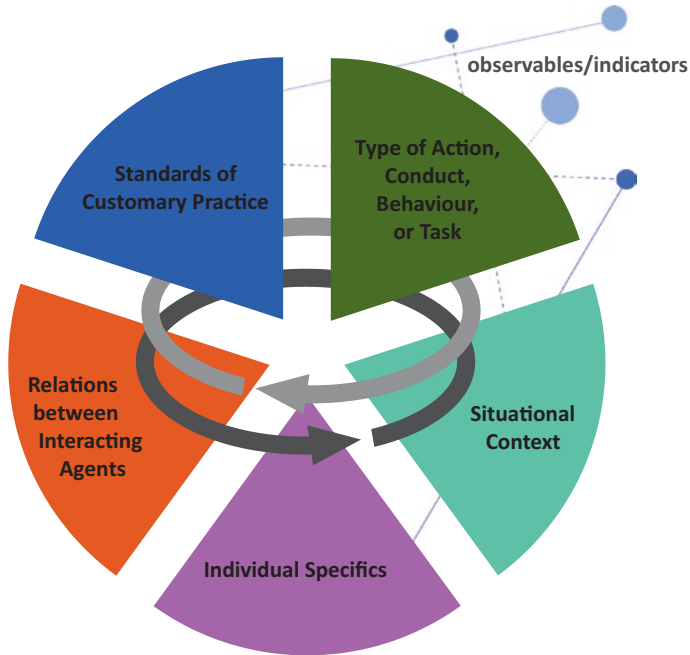


Fig. 4.1 Factors of social appropriateness. (Source: own representation)

arrows in Fig. 4.1—these factors and parameters of corresponding factor criteria influence what is considered socially appropriate in specific interactions.

EXAMPLE 1: THE FIVE FACTORS OF SOCIAL APPROPRIATENESS

Imagine the following situation: in a queue at a cash register, the person in front of you takes their groceries and leaves the store without paying. Perhaps you perceive this to be unusual and feel irritated (factors guiding construction and evaluation³ of appropriateness include: «Standards of

³ 'Factors guiding construction and evaluation' refers to the following questions: how can socially appropriate action be constructed? What are the factors used to judge the appropriateness of an action?

Customary Practice» of a **«Type of Action»** in the **«Situational Context»** of buying and selling). You may conclude that nothing strange is going on, especially if none of the cashiers or store security employees are protesting: You might suspect that the buyer is in fact an employee or the owner of the company (factors guiding construction and evaluation include: **«Individual Specifics»** and **«Relations between Interacting Agents»**) or perhaps that this business allows known and trusted (regular) customers to ‘open a tab’ (factors guiding construction and evaluation include: **«Individual Specifics»** and **«Standards of Customary Practice»**).

Example 1 shows that the mentioned factors in the construction and/or perception of the social appropriateness of an action or behaviour are not only **mutually interconnected**, but they are also relevant to **different extents**—in some situations, certain factors guide construction and/or evaluation more than others when judging the appropriateness of an action, whereby the action can be performed by oneself or by another agent. While the factors do influence one another, a single factor can be decisive in a particular case—the *FASA MODEL* reflects this by representing individual factors separately from one another, loosely arranged into the slices of a ‘pie’.

In figure 4.1, ‘observables’ refer to observable and potentially measurable features that can provide information to people or other context-sensitive (potentially technical) systems in specific circumstances—and to various degrees of ‘awareness’—about the concrete expressions of the factors. When information is interpreted directly from measurable and observable data by a person or technical system, the observables are called ‘indicators’. Thus, observables and indicators do not differ in terms of data, but in terms of the meaning assigned by an interpreter; see example 2. The information or meaning extracted by a person from purely observational data is typically not processed linearly in the sense that a particular factor *y* is inferred or reliably deduced from a particular observable *x* as an indicator for this factor, meaning is derived or assigned diffusely, sometimes preconsciously, and maybe most importantly: it is variable; especially with regard to when and why an observed parameter or behaviour is judged in terms of its social appropriateness. While human perception works this way, if we wanted to implement aspects of social appropriateness into technical systems, we would need to simplify these complexities significantly and tell the system how to interpret an observable as an indicator of social in/appropriateness.

Although certain typical associations can be made, different observables can indicate different factors to different levels of strength or clarity (indicated in Fig. 4.1 by the **different and sizes of the observables**). There is another level between the observables read as indicators and the factors of the model, namely

the ‘**factor criteria**’, which are essentially sub-factors representing possible features of the factors—from which further ‘conclusions’ about the factors can then be made. **A factor criterion, such as <time>, can be shared by multiple factors.** For example, the «Situational Context» factor is reflected in various factor criteria that are also associated with the other factors, since the latter also play a role in how the situation is perceived. At the same time, the situation is viewed as a factor in its own right, since the interacting agents must always decide which situation they are in. It is therefore not just a conglomerate of the parameters of the other factor criteria but a variable that influences behaviour in its own right.

Finally, it should be noted that it is possible for an observable to be perceptible ‘by itself’ and/or lie within the field of diffuse perception but not provide any information about a factor criterion or factor at all; in other words, as an indicator, it does not provide any insight into anything relating to social appropriateness. The assignment of factor criteria to factors and the subdivision of social appropriateness into the factors cited above is not ‘naturally’ predetermined. From an epistemic point of view, it is the result of linguistic and other structuring of the world by human perception and action. In practical terms, the division into factors and factor criteria is the result of the presented goal-oriented project work attempting to make the phenomenon of social appropriateness tangible. The list and classification into factors and factor criteria do not claim to be complete or exhaustive. Moreover, for some factors, it could undoubtedly be argued that they should not be listed separately but characterized as subordinate or superior to some other factor.

For example, if social appropriateness is understood as an **interaction phenomenon**, then, in a narrow sense, *only* an action or behaviour can be socially appropriate or inappropriate, and all other factors could be subsumed under the single factor «Type of Action, Conduct, Behaviour, or Task». It could also be argued that <time>, instead of being considered a factor criterion, should be considered a separate factor in some cases. In this regard, the factors represent our own attempt to reconstruct the ‘largest melting pots’ of what aspects most frequently and urgently influence the concept of social appropriateness, to the best of our knowledge. For example, «Type of Action, Conduct, Behaviour, or Task» is considered a factor precisely because only actions and behaviours can be socially appropriate when social appropriateness is viewed as an interaction phenomenon. On the other hand, specific moments in time are sometimes decisive for the construction and evaluation of social appropriateness, but in our opinion only to the extent that they provide information about other factors, such as the present situation, which is why <time> was listed as a factor criterion (rather than a factor). The factor criteria <time> and <intention> are associated with two different factors. As mentioned above, this association of <time> with both «Situational

Context» and «Type of Action, Conduct, Behaviour, or Task» and of «intention» with both «Type of Action, Conduct, Behaviour, or Task» and «Relations between Interacting Agents» illustrates the interconnections between the factor criteria and the factors, and therefore between the factors themselves. Analogously to the remarks mentioned above on factors and observables, additional intermediate features (i.e., factor **sub**criteria) might still need to be introduced to extend the factor criteria; if so, the possibility of being assigned to multiple factor criteria would apply equally to these subcriteria.

EXAMPLE 2: OBSERVABLES AND INDICATORS

An example that illustrates the difference between observables and indicators is the phosphate content of drinking water. Here, there is a measurable value (observable) that can be read as an indicator of drinking water quality. But this value only becomes an action-guiding criterion once the measured entity is interpreted as an indicator, of, for example, drinking water quality, and when in consequence certain thresholds are defined. Transferred to social appropriateness, for example, the volume of voices in a conversation can be observed. On its own, this does not mean anything and does not have an action-guiding function. Only when interpreted against the background of contextual factors does conversation volume serve as an indicator of what is socially appropriate in the specific situation and which factor criterion, or factor plays a role in the specific situation. To further illustrate the relation between observables and indicators, consider the following situation: when you ‘interpret’ the observable⁴ ‘clothing style and condition’ of a person as an indicator—without necessarily being aware of doing so—it might give you insight into the factor criteria «**social position**» (e.g., through professional attire) and «**individual characteristics**», such as the situational state of their physical, psychological, and cognitive constitution (in other words, it **signals/indicates** something to you about these factors).

For technical systems, some of the conditions that apply to the construction and evaluation of socially appropriate types of action, conduct, behaviour, or task are

⁴Simplified for purposes of illustration. The actual observables would for example be specific items of clothing, or even more fine-grained: specific sections of specific items of clothing. More precisely, one already ‘infers’ (in most cases implicitly and ‘automatically’) conclusions about the specific clothing style by identifying certain features whose situational meaning would first need to be explained to a technical system.

the same as in interpersonal interactions, and some are very different. The *FASA MODEL* allows aspects of human-machine interactions of all kinds to be inspected to determine which factors are being considered, which factors could or should be considered, and which factors cannot or should not be considered. The content of each factor is explained in more detail in the next few chapters, and a brief overview of the theories that play a key role in conceptualization and interpretation of social appropriateness is given. The lists of theories cited and discussed do not claim to be exhaustive but simply reflect the theories and concepts that were found to be highly relevant to the development and conceptualization of each factor during our research.

As an alternative approach to the phenomenon of social appropriateness, the factor criteria and observables/indicators of social appropriateness presented and discussed below in the context of the factors of the FASA model are interconnected and arranged into tree graphs in Chap. 5.

4.1 The «Type of Action, Conduct, Behaviour, or Task» factor

«Type of Action, Conduct, Behaviour, or Task»: A concrete action (more precisely: action sequence) or behaviour (more precisely: behaviour sequence) of a type of action, conduct, behaviour, or task can manifest either in execution or in omission mode; in other words, omissions are also actions. The factor «Type of Action, Conduct, Behaviour, or Task» is directly related to the other factors. For example, the typical interaction goals of a type of action, conduct, behaviour, or task can be established for specific types of situations, associated with corresponding typical behaviour expectations placed on the interacting agents and standard judgements of social appropriateness. But «Type of Action, Conduct, Behaviour, or Task» is also shaped by individual specifics in the conditions, preconceptions, goals, etc. of the interacting agents. According to psychological theory, the socially appropriate chain of action sequences in a given situation is stored in human memory as so-called ‘**scripts**’ (Abelson 1981; Nishida 2005; Schank 1975, 1982; Ziem 2008).

Possible factor criteria:

- **«Time»:** When is the behaviour taking place?
- **«Role identities»:** Who is performing the behaviour? Who is judging the behaviour?

- **«Intention»:** With what orientations—motivations, goals, intentions, focal points—and for what purposes is the behaviour/action taking place?
- **«Consequences»:** What consequences does, or could the interaction have?

Examples of questions and aspects that can be assigned to each factor criterion to support an analysis of the potential social (in)appropriateness of interaction situations and guide these situations are given in Table 4.1.

Possible observables/indicators:

- Duration of a conversation; form of a conversation with corresponding chronological requirements; duration and frequency of speech by specific interacting agents
- For both the acting and judging parties, aspects that play a role in judging appropriateness: gender, age, other group-related and/or (self) assigned individual features, such as ethnicity, social rank, attributed social capital (so-called ‘reputation’), other phenotypic features; codes of conduct
- Language use; gestures; facial expressions; action types and sequences, etc.
- Form and content of conversations; facial expressions; gestures; location-related and environmental features like buildings, physical space including any props or typical frame markers, e.g., for a theatrical play: stage, audience, curtain, lighting, tickets, programme announcing the performance, etc.

4.1.1 Theoretical perspectives on «Type of Action, Conduct, Behaviour, or Task»

4.1.1.1 Knowledge organisation and schemata

To acquire knowledge about what type of action, conduct, behaviour, or task is considered socially appropriate—and this applies equally to the other factors too—both people and technical systems need some general knowledge organisation system—a ‘database’ of knowledge or intuition about what is judged to be appropriate, by whom, and in what circumstances. For technical systems, knowledge modelling is for example provided by ontologies (Gruber 1995; e.g. Neches et al. 1991). For humans—according to psychological theory—knowledge is structured into so-called *scripts* for behavioural processes, and *schemas* for general knowledge organisation (see excursus on scripts and schemata).

Table 4.1 Type of Action, Conduct, Behaviour, or Task. (Source: own representation)

«Type of Action, Conduct, Behaviour, or Task»		
<p>«Time» When and how is the behaviour taking place?</p>	<p>How is it taking place? Phrased in language? Through facial expressions? Through gestures? How is it being executed?</p>	<p>If through language, for example: what type of conversation? A dialogue? Triadically? A multi-agent situation? What are the specifics of the interaction regarding taking turns, active listening, interventions, interruptions, relevance of grounds for interruption?</p>
	<p>Tact at the ‘right time’ (<i>kairos</i>^a): is it ‘too early’ or ‘too late’ to perform a certain type of action, conduct, behaviour, or task?</p>	
	<p>Is there a (repeatedly) recognizable structure of interactive, chronological sequences? What time intervals are typical and expected: in a conversation? For specific (‘scripted’) actions and types of behaviour?</p>	
<p>«Role identities» Who is performing the behaviour? Who is judging the behaviour? (closely related to «Individual Specifics» and «Relations between Interacting Agents»)</p>	<p>Attributions and performance of role identities</p>	<p>What social roles are attributed to the interacting agents? How can socially attributed roles and role identifications be recognized? To what extent do judgements of appropriateness change when the same behaviour is performed with different role identities? What needs to be reflected in specific roles (e.g., etiquette for aristocratic society/social club conversation rules, etc.)?</p>
	<p>Representation: are people acting or behaving directly on behalf of themselves, or are they being represented, or are they representing other people? How do customary practices differ between agencies/politicians and private individuals?</p>	
	<p>Role relationships: how do the role identities relate to one another? Resident or tourist? Citizen or state? Poor or rich? (see: «Relations between Interacting Agents»)</p>	<p>Expectations and expectations of expectations: to what extent does the expected acceptance (by both the acting party and the judging party) of a type of action, conduct, behaviour, or task change with self-perceived and externally attributed role identities (also by both parties)?</p>

(continued)

Table 4.1 (continued)

«Type of Action, Conduct, Behaviour, or Task»		
«Intention» With which orientations (motivations, goals, intentions, focal points) is the behaviour/action taking place?	Purpose of action: are there persuasive intentions? Cooperative intentions? Other intentions?	
	Goals and capabilities: how much willingness for cooperation and capability for cooperation is there? What other goals and capabilities are there?	Do the parties performing the action/behaviour and the parties judging the appropriateness of the action/behaviour have the ability to recognize emotions and situations?
«Consequences» What consequences does or could the interaction have?	Group-specific behavioural rules (see «Standards of Customary Practice»)	The degree of dissemination and observance of rules of conduct with a group of actors
	Visibility of consequences (see «Standards of Customary Practice»)	Visibility of consequences: are behavioural rules institutionally and normatively enshrined? Are they institutionally sanctioned?
	Existing power dynamics (see «Relations between Interacting Agents»)	What power dynamics are there between the interacting agents (cf. «Relations between Interacting Agents»)?

^aSee for example Liddell and Scott (1940).

Excursus on scripts and schemata

For human orientation about what action and/or behaviour sequences are appropriate, the psychological script theory (see e.g. Schank and Abelson 1977) postulates that sequences of specific action processes are stored in the human memory as schematic scripts. In general, SCRIPT THEORY can be viewed as part of SCHEMA THEORY. Schemata are postulated to be found in connection with knowledge organisation and in particular in human learning processes and the organisation of human memory. The conditions for retrieving a schema or activating it can relate to the internal and external entities present in a situation, or the concepts or event routines evoked by them. By referencing a previously established, known situation, in schema activation a frame is recognized that enables a person to ‘understand’ (‘interpret’) the situation, expand an existing frame of understanding, or even create a new understanding.

Schemata are the units of general knowledge organisation in which the accumulated knowledge of past experiences is stored, and they therefore also contain the regular features of what is considered socially appropriate. Cultural norms and expectations are stored both individually in schemata developed by individual subjects and collectively in shared schemata (cf. e.g. Bartlett 1932, 1958). Using schemata evoked by a situation, a person can adapt their modes of action and/or behaviour to make them situationally appropriate—but a person might err in this as well for several reasons, see e.g., sect. 4.1.1.2. One particular aspect of schematically organised memory is that the creation of a new schema by repeatedly experiencing sequences of a *token* (single instance) of its components means that specific details of individual situations are remembered less well once the new schema has been established. Thus, details are remembered better after a one-off occurrence of a certain situational process, whereas remembering a learned routine after repeated *iterations* makes remembering individual details harder (Hudson 1990): in humans, the denoting ‘register’—the schema—becomes more salient than the individual observables themselves after it has been established, whereas the individual observables play a more prominent role whenever no schema is available for the experience. New schemata can emerge and old ones can *fade*; in this sense, schemata are, loosely speaking, primarily a matter of habit.

Theories based on the concept of schema can be found in various scientific fields. One example is *FRAME THEORY* in linguistics (for an application of frame theoretical considerations with regard to social appropriateness in technical systems see e.g. Busse 2019). Frame theory is a theoretical approach to knowledge structures that transposes schemata to the context of purely linguistic organisation using the concept of *frames*: “Frames are conceptual units of knowledge that evoke linguistic expressions during language comprehension and which language users retrieve from their memory in order to grasp the meaning of a linguistic expression” (cf. Ziem 2008, p. 2, own translation). One noteworthy result of frame-theoretical linguistic studies is that “no criterion legitimizes a sharp separation between linguistic knowledge and world knowledge” (cf. Ziem 2008, p. 4, own translation). Approaches to possible formalizations of language-based schemata and world knowledge are for example offered by formal ontologies. The concept of Basic Formal Ontology for example bridges the gap between philosophical considerations about the possibility of formalizing ontologies and the actual implementation of these ontologies (Arp et al. 2015; Burkhardt and Smith 1993; Ceusters and Smith

2010). Example 3 deals with some of the problems of transferability of human world knowledge to technical systems.

EXAMPLE 3: CAN A ROBOT GIVE A TIP?

Consider an application scenario for a technical system based on the example of the linguistic *frame* of ‘tip’. For now, we are not interested in discussing whether it makes sense to build a technical system that pays tips or if this task should be left to a person rather than being operationalized. Instead, we ask *what would be necessary if operationalization were indeed our goal*.

In most adults from cultures with service industries, the word ‘tip’ evokes an association “in which there is a scenario (the restaurant), actors (guests, the waiter), some props (a table, chairs, cutlery, etc.), a script (ordering food, eating, paying), and much more” (cf. Ziem 2008, p. 2, own translation). In principle, “everyone knows under what circumstances and in which everyday situations tips should be given” (cf. Ziem 2008, p. 2, own translation). In practice, however, this does not hold true for everyone, and especially not for technical systems, for which such forms of everyday knowledge are difficult to make accessible. How can we—as a basis for socially appropriate action and behaviour sequences with regard to *tipping*

- a) teach a technical system the concept of tipping?
- b) equip this system with the corresponding socially appropriate action and behaviour sequences?
- c) expand the system’s sensors to recognize a *tip transaction* situation?
- d) teach the system *when and where* tipping is appropriate through observables?
- e) teach the system cultural differences in tipping habits?

Potential answers to these questions imply that declarative (*knowing-that*), as well as procedural (*knowing-how*) knowledge is necessary for socially appropriate behaviour (cf. Nishida 2005, p. 408 referring to Tulving 1985; Zola-Morgan and Squire 1990).

Regarding schemata and scripts for human orientation regarding the appropriateness of types of action, conduct, behaviour, or tasks Taylor and Crocker (1981) suggest differentiating between at least **five different types of schemata**

(regarding the problem that infinitely many nested schemata are principally possible, see also Busse 2019), namely:

(a) Person schemas, which are knowledge about different types of people, including their personality traits; (b) self schemas, which contain knowledge about themselves; (c) role schemas, which represent knowledge about social rules; (d) event schemas or scripts, which are information about the appropriate sequence of events in common situations; and (e) content-free schemas, which are information about processing rules. (following Nishida 2005, p. 405 referring to Taylor and Crocker 1981)

Hiroko Nishida (2005) in his overview on schematic knowledge organisation⁵ names a total of eight “**primary social interaction schemas**” (PSI SCHEMAS, cf. Nishida 2005, p. 407):

Fact-and-concept schemas	for example “Tokyo is the capital of Japan” or “Bicycles are those vehicles that have two wheels, a seat, and handlebars” (ibid.)
Person schemas	for example “John is neurotic”, “Taro is shy”, “Mary is easy-going” (ibid.)
Self schemas	for example <i>I am generous</i> or <i>Other people perceive me as generous</i> .
Role schemas	contain social stereotypes such as <i>Women work hard</i> , <i>Men cook well</i> , or <i>Poor people have good manners</i>
Context schemas	contain rules of appropriateness , activated before any procedure schemas or problem-solving schemas to specify their appropriateness, e.g., <i>It is usually appropriate to sit on a chair or A chair on a theatre stage is usually only for actors to sit on, not spectators</i>
Procedure schemas (scripts)	specify the appropriate sequence of actions in typical situations, such as <i>After paying for my ticket, I enter the cinema room, I can buy popcorn in between.</i>

⁵ ‘Schemata’ is used throughout the text at hand to refer to all schemata, ‘schemas’ is used as the technical term in psychological theory.

Strategy schemas

contain problem-solving schemas that depend for example on the level of expertise of the acting parties, e.g., *If I think that I might be late, then I will hurry or if a person is tense, they get a massage, if I want to get better at playing the piano, I must practise*

Emotion schemas

contain information about affective states and possible feelings, such as *the death of a friend makes you sad*.⁶

Each of the various schemas has more general and more specific sub-schemas, which may overlap in some regards. Furthermore, within each schema, there can be links to other schemas; for example, context schemas contain connections to role or procedure schemas: “For example, when ‘being at a dentist’s office’ (a context schema) is selected, the role and procedure schemas that are appropriate in the context are activated.” (Nishida 2005, p. 410).

Regarding «Type of Action, Conduct, Behaviour, or Task», we are especially interested in procedure and context schemas in the context of social appropriateness. Procedure schemas are also captured by the concept of ‘scripts’. Scripts are a form of (sub-)schema in which event sequences are stored. Script theory can be traced back to the psychological research by Schank and Abelson (Schank 1975; Schank and Abelson 1977) and includes a dimension concerning culture (“*cultural scripts*”), which is relevant to the factor «Situational Context» and its cultural dependence. It also includes a distinction between *mindful* (script-oriented) and *mindless behaviours* (automatically performed *habits*) (Abelson 1981). Similar theories also motivated by schema theory are found among the cognitive sciences with Marvin Minsky (1974, 1977) and Lawrence Barsalou (1992), as well as in the above-mentioned linguistics with Charles John Fillmore (1975, 1977, 1982), where additional concepts such as linguistic *frames* are introduced. Regarding scripts organised in memory, Roger Schank also speaks of *Memory Organisation Packets* (MOPs) in which the individual scenes used to achieve an interactional goal or master a situational context are grouped together in a memory organisation system as packets. The scene or script of ‘paying’ (form: [order,

⁶For various other definitions and areas of application of the concept of schema in psychology, see also: Kiss et al. (1993); Mandler (1984); Rumelhart (1980); Schank and Abelson (1977); Taylor and Crocker (1981); Thorndyke (1984); etc. (cf. Nishida 2005).

receive, pay]) is for example encountered in various contexts such as cash register, restaurant, tips, bribes, and the relation between ‘paying’ and the respective context is organised in terms of MOPs (cf. for example Schank 1999, p. 123–136).

Since schema and script theories are good at describing information perception and processing in general, they do not only explain human communication and interaction processes but have also been used in the context of computer-based perception and memory simulation. For example, Turner (1994) named three schemas considered to be important for problem-solving by a *schema-based reasoner* such as an AI:

- (a) procedural schemas, which contain information about steps to take or hierarchical plans; (b) contextual schemas, which are information about the situation or appropriate setting of behavioural parameters; and (c) strategic schemas, which are knowledge about problem-solving strategies. (Nishida 2005, p. 405)

That is to say, psychological schema theories have explanatory power for the creation and change processes of social appropriateness and its judgement, while also providing a direct bridge to the technical communicability of these processes—with some limitations, as we will see below. Also notice that the quote refers to “information about the situation” which points at its relevance for the factor «Situational Context» (sect. 4.2).⁷

⁷According to Nishida, Turner postulates the following sequence in the production of behaviour: “(a) When individuals are in a specific situation, they try to recognize whether they know the situation by retrieving one or more context schemas from memory that may represent the current situation [i.e., recognition of observables, conclusions about observable fields and contexts by reading observables as indicators, remark by the authors]; (b) when an appropriate context schema, which represents similar interaction situations, is found, the context schema subsequently suggests a goal to pursue (through an attention-focusing function of the context schema); (c) when a goal is selected, the context schema looks for a strategy usually useful in situations of this sort; (d) the context schema then suggests a procedure schema with which to achieve the goal, using the selected strategy schema (i.e., the procedure schema is retrieved for taking specific actions; in other words, the procedure schema specifies steps to take, hierarchical plans, or behavioural rules); and (e) the application of the procedure schema causes an individual to take some action, such as asking questions of the other interactant, which further causes the individual to specify the current context more clearly; in other words, to find a context schema that is more specific for the situation. When the more specific context schema is applied, the schema

4.1.1.2 Social Information Processing model and Frame Selection model

Once certain perception templates are available—phrased more generally, once entities can be distinguished from one another, and, named if necessary—information can be processed. To illustrate the chronological process sequence in human SOCIAL INFORMATION PROCESSING, the model of SOCIAL-COGNITIVE INFORMATION PROCESSING (Crick and Dodge 1994; Dodge 1986; Dodge and Pettit 2003; Gifford-Smith and Rabiner 2005), which bears at least some resemblance to schema and script theories can be used.⁸

The SOCIAL INFORMATION PROCESSING MODEL as reproduced after Döpfner 1989 in Fig. 4.2 distinguishes eight phases of information processing. The model postulates that these phases occur in every social situation:

Phases can overlap and individual phases can “be skipped and made up later” (Döpfner 1989, p. 4, own translation). Information processing and therefore the execution of actions can be impaired in any phase, for example if

- (a) relevant signs from the social situation are not perceived or are misinterpreted;
- (b) possible alternative courses of action are not taken into consideration;
- (c) consequences of action are ignored or evaluated incorrectly;
- (d) alternative courses of action are associated with unrealistic results or competence expectations;
- (e) decisions are made impulsively;
- (f) action is insufficiently planned;
- (g) the actual consequences of an action are incorrectly registered, causally misattributed, or not processed. (cf. Döpfner 1989, p. 5, own translation)

If there are disruptions in information processing, for example if “social cues are misinterpreted as being threatening” (cf. *ibid.*, own translation), the subsequent problem-solving process can lead to “wrong decisions” (cf. *ibid.*, own translation), which can of course also influence the appropriateness of a certain type of action, conduct, behaviour, or task (see example 4).

further finds appropriate strategy and procedure schemas to apply to the new context. The more specific the current context schema is, the better the chance that more-specific strategy and procedure schemas will be suggested.” (Nishida 2005, p. 407 referring to Turner 1994).

⁸For an overview, cf. e.g. Dodge and Rabiner (2004).

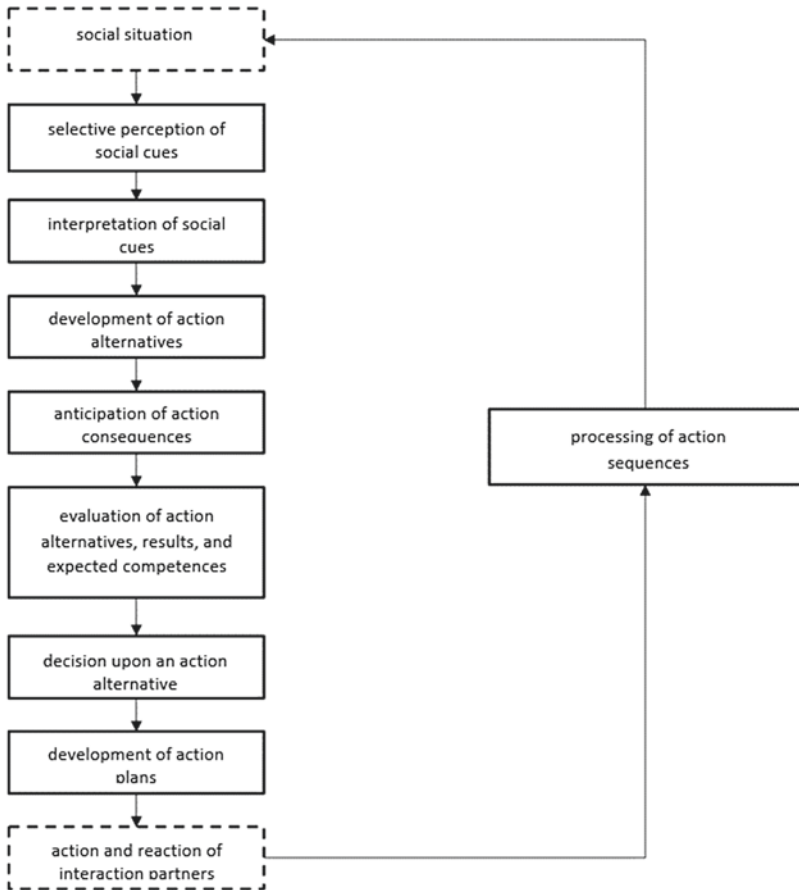


Fig. 4.2 Social information processing model. (Adapted from Döpfner 1989, p. 4)

EXAMPLE 4: DISRUPTIONS IN INFORMATION PROCESSING

Two distant acquaintances run into each other by chance in the city. They see and greet each other, but one of them appears to be in a hurry. This is where the first mistake might be made. The hurrying person may display (social) cues of their haste, for example by looking fidgety or checking the time. These cues may or may not be perceived by the other person.

If this information is not perceived, the interaction partner will not adapt their behaviour to accommodate the other person's haste and may therefore hold them up inappropriately instead of showing polite consideration. Mistakes can also be made when interpreting social cues. The non-hurrying person might for example simply assume that the other is a generally fidgety person, which could lead to the same inappropriate behaviour as described above. When generating possibilities for action, a mistake could be committed by failing to generate the right alternative, e.g., simply asking whether the other person is in a hurry. Likewise, a mistake might be made when anticipating the consequences of an action, leading them to select the wrong alternative. The two steps of evaluating each alternative and deciding on one of them go hand in hand. For example, asking whether the other person is in a hurry might be internally judged an invasion of privacy and not selected for this reason. When developing an action plan, mistakes can once again occur, for example if the action plan is disrupted by a distraction (someone else walks by and is briefly greeted) and the action is not carried out. Finally, mistakes can be made when processing the consequences of an action by misinterpreting the behaviour of the other person. Any of these mistakes would cause one person to neglect consideration of the other's haste, causing the conversation to be unnecessarily prolonged. If the other person then says goodbye, perhaps abruptly in a manner that is itself socially inappropriate, the first might ponder what they did wrong, but they might also simply conclude that the other person was having a bad day. If so, nothing is learned from the situation and the behaviour is not corrected.

The step-based model of information processing was originally understood as analogous to technical information processing (Dodge 1986), but researchers later turned to see human perception processes and decisions, unlike machine processes, as inevitably influenced by their emotional states (Dodge 2010, [1991]).⁹ This difference should not be neglected when transposing structures of human information processing over to technical systems in the context of socially appropriate types of action, conduct, behaviour, or task.

⁹This contrasts with models of human decision-making based on purely economic principles, such as RATIONAL CHOICE THEORY and GAME THEORY (sect. 4.5.1.2).

The sociological MODEL OF FRAME SELECTION (MFS) developed by Clemens Kroneberg based on work by German sociologist Hartmut Esser as shown in Fig. 4.3 explains (a) how actors interpret situations, (b) which action scripts they activate, and (c) which actions they perform (Kroneberg 2010, 2011). Human actors make reflective choices or base their decisions on strong emotions, normative convictions, or unquestioned routines that ignore objectively relevant alternatives and incentives. The MFS distinguishes between three substantive selections: the selection of a **frame** (situation definition), a **script** (action programme for defined situations), and an **action alternative** (intended action). The MFS highlights the importance of situation definitions through frame and script selection, it is therefore also relevant to the «Situational Context» factor (see sect. 4.2.1.1). Nonetheless, the model is introduced at this point to illustrate the similarities and differences to the information processing model as mentioned above. Again, errors can occur at each step leading to a socially inappropriate action being selected to perform in the end, but there is more focus on the interpretation of a situation ('frame selection'), whereas the focus in the model above is on social cues that would be misinterpreted. To give an example related to social appropriateness, a situation characterised by the fact that everyone is wearing black clothes and is crying could be read as a funeral, but it could also be a theatre

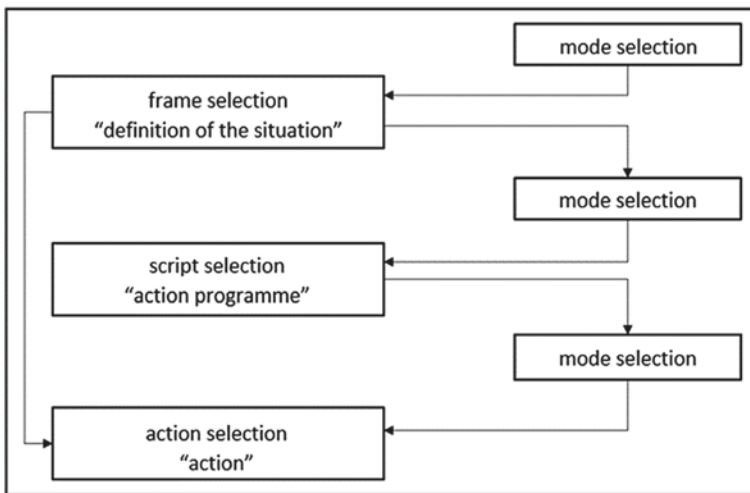


Fig. 4.3 The types of selection considered in the frame selection model (Kroneberg 2010, p. 133)

play in which a similar scene is being re-enacted, or even a meeting of people who belong to a subculture and are experimenting with some type of ritualized behaviour. However, the frame [funeral] is very different from the frame [theatrical play] or [experimental self-experience] and contains completely different behavioural scripts.

Errors can also occur within a script, for example because the expected behavioural script is not fully known to a person (e.g., having a different cultural background). This can in turn lead to an incorrect choice of action alternative. Socially appropriate behaviour is only achieved when frame selection is carried out correctly—i.e., often, according to the other interacting agents' selection.

4.1.2 «Type of Action, Conduct, Behaviour, or Task» of sociosensitive/socioactive artificial assistants

Especially when considering the difference between declarative and procedural knowledge, one question that plays a key role for the factor «Type of Action, Conduct, Behaviour, or Task» in the context of socioactive and socioactive systems is the extent to which we attribute agency or 'self-chosen' behaviour options to machines, as well as the ability to 'know things' in general. There are various theoretical understandings of how behaviour, action, and knowledge relate to one another, for example in sociology (Vester 2009). Although the text at hand does not consistently differentiate between action and behaviour, our understanding of the distinction roughly follows that of sociologist Walter Bühl (1982): "Humans can act and recognize things, but only within the framework of a biologically pre-structured behaviour space, and not outside of it" (cf. *ibid.*, p. 172, own translation). This also corresponds to **one possible** philosophical approach to the action and behaviour complex, which namely states that human behaviour cannot be traced back to reasons and motives, in essence: it does not unfold in a purpose-oriented manner, rather it represents organic or motor reactions to stimuli or other basic biological conditions (for a classification of this key philosophical terminology, see for example Jantschek 2018 for 'behaviour' and Lorenz 2008 for 'action'). As an example to illustrate this and demonstrate the further differentiation of behaviour and action, as well as action into social action:

Behaviour is when someone is sweating. Action is when a person who is sweating seeks some type of refreshment, since they associate their action (fetching a drink or fanning some cool air) with a subjective meaning. If this person now decides to use a deodorant (to spare their fellow human beings the unpleasant consequences of perspiration), we can describe this as a social action. (Vester 2009, p. 46)

Max Weber fundamentally defines action as a human “behaviour—be it overt or covert, omission or acquiescence” with which the actor “attaches a subjective meaning” (cf. Weber 1978, p. 4). In the context of social appropriateness, we are in particular interested in further differentiating this into the concept of **social action**: “Action is ‘social’ insofar as its subjective meaning takes account of the behaviour of others and is thereby oriented in its course.” (cf. *ibid.*).

Regarding technical systems, we might therefore ask to what extent these systems (can) relate to the behaviour/actions of other parties (humans, other technical systems) exhibiting actions/behaviours. Algorithms trading on the stock exchange could be mentioned as one example of technical systems interacting in such a way they react to each others’ actions/behaviour. For various reasons, one of them being the wish to be able to explain interactions in which technical systems take other system’s behaviour into account, Rahwan et al. (2019) call for a **science of machine behaviour**. The authors propose a few suggestions about what an approach to a science of machine behaviour might look like. But even if we study machine behaviour and attest that they do to some extent take into account the behaviour of others and orientate their own behaviour towards it, they still do not attach subjective meaning to this in the sense humans do; technical systems only ever ‘act’ at the process level, i.e. they ultimately only process, but do not form their own wishes, normative attitudes, or desires. If we understand behaviour and action as the experience and expression of biological conditions and as the means-based realization of a purpose, technical systems cannot be said to ‘act’ or ‘behave’ at all—although *functional equivalents* of intrinsic motivation may be implemented (cf. for a consideration about functional equivalents of emotion components in technical systems also Bellon 2022).

Even though technical systems cannot be ascribed full agency and do not really act, let alone socially, in the sense human agents act, ‘just processing’ still has effects on humans. This can be seen for example in technically mediated *nudging*¹⁰ (Thaler and Sunstein 2009), as well as in technically mediated (dis)information influencing people’s voting behaviour. As for the question of agency and responsibility, in the end, ultimate responsibility lies with the people who design, install, and maintain the technical system, or, follow its instructions or suggestions.

With regard to technical artificial assistants and types of action, conduct, behaviour, or task, artificial assistants can support action to facilitate success where it

¹⁰NUDGE OR NUDGING THEORY postulates that individuals or groups of people can be prompted to engage in certain modes of action and behaviour through positive *reinforcement* and indirect suggestions.

would otherwise not be possible or would be less successful. But technical assistance can also make independent action impossible by taking over processes that must be performed in person as a prerequisite for the cultivation of competencies (deskilling, see also chapter 7).

4.2 The «Situational Context» factor

The participants of simple interaction systems can only start, shape, finish, or continue communication if they develop a shared definition of the situation. (Preyer 2012, p. 151, own translation)

The «Situational Context» of a specific action or behaviour is characterized by (c.f. March 1994; Busse 2019):

- a specific place and time,
- a specific spatial structure,
- a degree of formality (intimate, familiar, private, semi-private, public, etc.),
- a degree of seriousness (serious, ironic, playful, etc.),
- a definition of typical roles,
- behaviour and/or action requirements,
- status requirements between the interacting agents,
- possible consequences,
- relationships governing the influence of various social rule systems,
- the interacting agents' definitions of the situation,
- and other aspects not yet considered here.

Possible factor criteria:

- **«Place and time»:** Where and when is the behaviour taking place?
- **«Framing»:** 'As what' is the behaviour taking place? For example: 'as' a job application? 'As' a theatrical play?
- **«Media-based and performative mediation»:** How is the behaviour being presented and performed?
- **«Participants»:** Who is interacting? Who is the behaviour targeting, who is judging it?

Examples of questions and aspects that can be assigned to each factor criterion to support an analysis of the potential social (in)appropriateness of interaction situations and guide these situations are given in Table 4.2.

Table 4.2 Situational Context. (Source: own representation)

<<Situational Context>>		
<Place and time> Where and when is the behaviour taking place?	<Place> : Where is the behaviour taking place?	What culture (religion, prevailing narratives and types of narratives, narrative styles, ideals, norms)? What subculture: are there any rules deviating from average norms for the purpose of delimitation? Public or private? City/country-side? Neighbourhood? Company-specific? Micro/macro level of social action?
	<Time> : The appropriateness of a behaviour depends on the time at which the action/ behaviour is performed (cf. action and behaviour criterion <time>)	What chronological sequences play an important role in the situation?
<Framing> ‘As what’ is the behaviour taking place? For example: ‘as’ a job application? ‘As’ a theatrical play?	Is there a schematically accessible, i.e., (repeatedly) recognizable, structure of interactive sequences?	
	Are there recognizable atmospheric features or indications of how atmospheric ‘data’ are (could be) perceived by the interacting agents?	
	Ritual/ritualized action? For example: opening ceremony? Dance? Dining/drinking ceremony? Wedding? Funeral? What kind of ritual manners are being performed? For example: table manners Is there public/private expression of respect/sanctioning?	
	Degree of seriousness? ‘As’ a game, or ‘seriously’? ‘As’ commentary, ‘as art’ (see <Media-based and performative mediation>)? As a conflict?	
	What ‘guiding differences’ (“ <i>Leitdifferenz</i> ”, see Luhmann 1984) must be observed as observation criteria, e.g., true/false or refutable/not refutable for a situation unfolding ‘as science’?	
	<Media-based and performative mediation> How is the behaviour being presented and performed?	Where: On media like the internet, blogs, social media, printed media, TV, cinema, advertising, text messages, picture messages Genre: ‘as what’: news, satire, theatre, reality TV, acting, staging, documentation, documentary fiction, science, etc

(continued)

Table 4.2 (continued)

«Situational Context»	
	Do specific rules/practices apply, such as e-conversation norms in emails, on specific forums (different norms or creation of independent rules, for example in online multiplayer games or on the 4chan image board), in presentations, in art, is there specific language behaviour?
	To what extent does the representation or performance of a situation fulfil/disappoint medium-specific expectations? To what extent was the situation created by the representation or performance of a social behaviour?
«Participants» Who is interacting? Who does the behaviour target, who is judging it? (Partly overlaps with the «Relations between Interacting Agents» factor)	Animals, objects, people, robots?
	Age, gender, height, other physical or other properties? Norm variant?
	Social roles? Attributed properties?
	Individuals, collectives? In-group/Out-group? How many participants? Does that make a difference?
	Are the participants present at the same or at different times? How do they feel?

Possible observables:

- Living space; building architecture (*high-* and *low-ranking buildings*, see Vanucci 2014); landscape
- Duration of a conversation; form and content of a conversation, etc.
- Typical requisites/frame markers, e.g., for a theatrical play: stage, audience, curtain, lighting, tickets, programme announcing the performance, etc.
- Media type and platform
- Markings of participants: gender, age, other group-related and individual features, such as attributed ethnicity, social rank, social capital, attributed social capital (so-called ‘reputation’), other phenotypic attributes; codes of conduct
- Participant’s attitudes and emotions

The «Situational Context» reflects various factor criteria that are associated with other factors, since the latter also play a role in how the situation is perceived. Nonetheless, the situation is itself viewed as a factor in its own right, since the interacting agents must always decide which situation they are in. It is therefore not just a conglomerate of the parameters of the other factor criteria but a variable that significantly influences behaviour in its own right.

4.2.1 Theoretical perspectives on the «Situational Context»

Below, we will consider some of the factor criteria individually as examples, such as the definitions of the situation by the participants and the culture surrounding people (as a sub-criterion of <place>) from the perspective of various theories. These examples illustrate aspects that could influence the construction and evaluation of what is perceived as socially appropriate in terms of the «Situational Context» factor. How or why a situation is experienced or defined one way or another depends on many different parameters and can shift abruptly or over time.

4.2.1.1 Defining a situation

The actors' situation definitions play a crucial role as framings for what will be perceived as appropriate social behaviour. If interacting agents have differing definitions of the situation, they will most likely consider different behaviours appropriate. To conceptualise situation definitions, we can draw from theories in sociology, psychology, philosophy and anthropology. A person's situation definition, no matter in what disciplinary terms you will describe it, will most likely depend on other factor criteria as mentioned in Table 4.2, such as the time and place of an interaction, thus, these criteria are mentioned throughout disciplines.

Sociologist Erving Goffman proposed the notions of **front stage** and **backstage** to grasp the fact that there are different standards of what is considered socially appropriate behaviour in areas with different levels of public visibility: for a waiter, for example, there are different standards of behaviour in the 'front stage' of a restaurant than 'backstage' in the break room (Goffman 2017). Here, the level of public visibility is the decisive criterion for defining the situation. This relates to the degree of formality mentioned above (sect. 4.2). Clemens Kroneberg, building, among others (see below), upon German sociologist Hartmut Esser and social phenomenologist Alfred Schütz, on the other hand emphasises on the joint action and cultural dimensions included in defining what situation interacting agents are in:

Every action is guided by a subjective *definition of the situation* (Goffman 1974; Parsons 1937: 76f.; Thomas and Znaniecki 1927: 68ff.): to be able to act, human actors must develop an interpretation of their situation. To do this, they often appeal to socially shared frames of reference that reflect the culturally defined meaning of typical situations and indicate their situational relevance through certain perceptible objects. (Kroneberg 2011, p. 12, own translation)

The three mentioned notions of “socially shared frames of reference”, their culturally defined meaning, and “certain perceptible objects” play significant roles in all five factors of our model: social behaviour—as we understand it—is embedded **in socially shared frames of references at all times and in all places** (relevant to the factors «Relations between Interacting Agents», «Standards of Customary Practice», and «Individual Specifics», at the very least). The concept of socially shared frames of reference reflects the fact that at least one acting person and one judging person is necessary to form a dyadic relationship and therefore a **social context**. If acting and judging are performed by one and the same person, i.e., if a person reflects upon their own actions, we may assume that the voice judging the appropriateness of an action or behaviour only exists because it has been brought to the person’s attention from some external source somewhere, sometime. The significance of **cultural meanings for ‘typical situations’** is further expounded below (sect. 4.2.1.2). Kroneberg’s reference to perceptible objects corresponds to our approach of reading observables as indicators to approach the question of the recognizability of ‘typical situations’.

Kroneberg continues with explanations referring to, among others and in the terminology of our model, aspects of «Standards of Customary Practice» and «Individual Specifics» and contrasts the idea that human behaviour is always based on rational choices (see also RATIONAL CHOICE THEORY, sect. 4.5.1.2), but is rather based on unconscious and more automated decision-making:

Secondly, human behaviour is characterized by *variable rationality* (Schütz and Luckmann 1979; Weber 1980). Although the ability to reflect and anticipate is a characteristic feature of human actors, humans do not always need to reflect on current situations and the actions to be taken. Thus, actors frequently follow their first impressions and the first behavioural dispositions that are activated. Only sometimes and in some respects do they make elaborate decisions with a systematic consideration and weighing of specific pieces of information and the expected consequences. (Kroneberg 2011, p. 12)

The aspect of «Individual Specifics» and Relations between «Interacting Agents» is further explored in the work of Weber, Kopelman and Messick (2004) and includes various factors that will only be touched upon here and will be discussed more elaborately under the factor «Individual Specifics», particularly with regard to social dilemmas (see sect. 4.3.1.3), which again can be grasped as situations. For our current purposes, social dilemmas can initially be understood as situations whose complexity is especially difficult to model technically but which can be described in terms of game theory and are therefore straightforward to

formalize (see sect. 4.5.1.2). The authors' following question aims at grasping some aspects that may come up in the process of decision-making and shows the extent to which «Individual Specifics», «Type of Action, Conduct and Behaviour», «Standards of Customary Practice» and «Situational Context» are mutually interrelated:

What does a person like me (identity) do (rules) in a situation like this (recognition)? (Weber et al. 2004, p. 282).

The authors point to the relevance of the *situation definition* of the person trying to answer this question:

The definition of the situation informs the person about the norms, expectations, rules, learned behaviours, skills, and possible strategies that are relevant. (...) The definition of the situation suggests a choice set. Choosing among the options, we contend, is a rule-directed exercise. (Weber et al. 2004, p. 285)

If a person decides what situation they are in and what role they play in it, they will then 'know' what to do, or have a "choice set" of possibly appropriate or inappropriate behaviour at hand. Information about what situation one finds themselves in is, on the other hand, provided by the ways in which the above-mentioned factor criteria manifest, and, of course, again by a person's knowledge organisation and personal and cultural background.

The process of ascribing certain values according to cultural standards is explored in detail below (sect. 4.2.1.2). The fact that certain objects can be involved in defining a situation is worth mentioning here again regarding the fact that observables become indicators to possible action and behaviour according to certain and intertwined characteristics of a person and their environment.

One approach to **formalizing the definition of a situation** with regard to the relations of a person and their environment, again taking into account phenomenological aspects of perception, is offered by the FIELD THEORY and TOPOLOGICAL PSYCHOLOGY of Gestalt psychologist Kurt Lewin (Lewin 1926, 1936, 2012), who proposed the following formula to explain behaviour:

A person's **behaviour** (B) is a function of the relations between this **person** (P) and their **environment** (E). The relations between person and environment constitutes both the persons' **living space** (L) and **situation** (S), so that

$B = f(PE)$; read: behaviour is a function of the relations between a person and their environment, whereby (PE) includes the living space (L) of a person as well as their situation (S) and definition of their situation; behaviour therefore also is a function of a person's living space and of their situation and situation definition.

Lewin emphasizes that a person's living space and situational context depend on both "the nature of the surrounding field" (Umwelt; environment) (Lewin 1936, 33) and the persons' current needs and specifics, both being structured by several principles. Speaking with our model terminology, Lewin emphasizes on the fact that observables become indicators according to both the object's and the perceiver's specifics. Objects in the environment according to him and in his terminology have a certain '*Aufforderungscharakter*', literally meaning they 'ask' a person to do something, later translated as **valences** by Lewin himself. A hungry person, for example will perceive food as stimulating and will organise their behaviour accordingly, a landscape will appear very different to a person depending on context, for example, in a **war landscape** as perceived by a soldier there may appear **battle things and peace-things** and single objects might change their 'Gestalt' from appearing as one to appearing as the other (Lewin 2009): in times of peace, it might seem almost barbaric to burn doors and furniture; the '*Aufforderungscharakter*' of burning is not contained within them; but in times of war, it might seem plausible and natural to exploit doors and furniture as heat-supplying objects or flaming barricades. Thus, this type of usage appears to a human perceiver as a 'stimulating character' or valence *within* the object. Nevertheless, it is simultaneously the living being's needs and definition of the situation that generate—in combination with the principal characteristics of that object—these 'stimulating characters'/valences in the first place.

This differs from the related term of *affordances* coined by the psychologist James Gibson (Gibson 1982, 2014) in that an *Aufforderungscharakter* also depends on the person perceiving it, whereas the concept of *affordances* focuses more on the object's properties inviting a person to act (e.g., through its design). The difference between these theories can be roughly summarized as follows: "A letterbox only has an '*Aufforderungscharakter*' for the observer if the observer wishes to post a letter; whereas the affordance [of the object, remark by the authors] still exists if there is no person present who wishes to post a letter." (cf. Hegenbart 2019, p. 58, own translation). Of course, it should be noted that design decisions are not independent of people and reflect human perceptions and basic prerequisites: the opening of a letterbox is placed at around hand height

because it is oriented towards the physical specifications of adult humans and exists because of the human need to post objects. But there is still a difference between the notion of *affordance*, being a stable property of the object—as in: the letterbox provides the principal possibility to post letters or more generally, insert things in it—and the notion of *Aufforderungscharakter*—as in: the letterbox stimulates a person to send a letter to communicate with someone, or, according to their needs and wishes, to insert a firecracker and let the box explode.

Now regarding social appropriateness, according to both Gibson and Lewin, it is not only *objects* that offer up valences or affordances; humans and other living beings also offer behavioural options and (perceived and/or displayed) stimuli to one another:

behaviour affords behaviour, and the whole subject matter of psychology and of the social sciences can be thought of as an elaboration of this basic fact. Sexual behaviour, nurturing behaviour, fighting behaviour, cooperative behaviour, economic behaviour, political behaviour—all depend on the perceiving of what another person or other persons afford, or sometimes on the misperceiving of it. (Gibson 1979/2014, p. 58).

With regard to human-machine interactions between humans and technical artificial assistants, it is especially important to note that machine behaviour—in the broadest sense, referring not only to the behaviour of robots but also to that of algorithms—‘offers’ specific types of behaviour to humans. We do not even need to go so far as to invoke the human tendency to perceive technical objects as beings capable of agency. Even where human actors do not ascribe any conscious or deliberate decision-making, agency or organic impulses to technical objects, they nonetheless potentially perceive them as (active) parts of their living space—and therefore also as elements relevant to the definition of their situation. A technical object’s behaviour may therefore be viewed as appropriate or inappropriate in the context of a particular situation definition, but it may also *invite a person to define a situation a certain way* or even act a certain way. Since technical objects may directly influence human behaviour by participating in the human living space they can be developed accordingly. Regarding technology development, we must bear in mind which goals are being pursued and which ethical considerations should be given attention. With regard to the factor «Type of Action, Conduct, Behaviour, or Task», we can perhaps add here that it might be desirable to have a science of machine behaviour (as for example Rahwan et al. 2019 ask for) exploring machines’ behaviour, but also taking into account the human motivations present in the engineering process and in the implementation of the technical object’s into societal, personal, interactional, or other structures.

4.2.1.2 Cultural schemas, cultural intelligence and culture dimension theories for situation understanding

There are many definitions of culture, and they can range from visible elements, such as the music styles of a group of people (or ‘pop-culture’), to non-visible elements, such as certain behavioral tendencies in a group. Some researchers take a view that includes both elements such as Stella Ting-Toomey’s conceptualization of culture: “Culture is like an iceberg: the deeper layers (e.g. traditions, beliefs, values) are hidden from our view; we only see and hear the uppermost layer of cultural artifacts (e.g. fashion, trends) and of verbal and non-verbal symbols” [50]. She then explains that to truly understand the nature of a culture, we must be able to match the upper layers with the lower layers. (Mascarenhas et al. 2016, p.933 quoting Ting-Toomey 1999)

When talking about «Situational Context», the factor criterion **«place» with a special focus on culture** deserves special attention. The cultural context in which technical systems will be used is relevant to their development. Technical systems’ behaviour and design may have to be adjusted to cultural standards to avoid rejection of the technical system, to make technically collected study results comparable, or for some other purpose. For example, researchers try to find out about preferred spatial distance of agents in human–machine interaction (‘proxemics’) in different cultures (Mead et al. 2013; see e.g., Rios-Martinez et al. 2015) to enhance user satisfaction with robotic systems.

In this chapter, culture is singled out and discussed as an example of the criterion «place» influencing a person’s situation understanding and related appropriateness judgements, although it is only *one aspect* of the parameters along the lines of which a situation will be perceived, defined, understood, or evaluated. Culture specific aspects of the construction and judgement of situationally appropriate behaviour will be presented from a mostly psychological perspective. Social psychologists such as quoted above assume that from observing or measuring, for example through surveys, the ‘uppermost layers’ of a culture, it may be possible to infer participants’ beliefs and values, which will influence what will be perceived as appropriate behaviour in certain situations—and what will be perceived as a what kind of situation in general.

According to a **theory of cultural schemas** (cf. for example Nishida 2005) a person stores a cultural schema (sect. 4.1.1.1) for known situations that will, among other information, such as what items will be likely to be present, provide knowledge about which behaviours are socially appropriate or inappropriate in this situation:

Cultural schemas for social interactions are cognitive structures that contain knowledge for face-to-face interactions in one's cultural environment. (Nishida 2005, S. 403)

The schema of a child's birthday party, for example, is different for a Mexican child than for a European child. The observable of *piñata* is only part of the birthday party schema if the child has previously seen, been told, or otherwise experienced the battering of a cardboard figure filled with candy. Changes in psychological schemas are for example described by Jean Piaget (1970, 2002 [1928]) with the concepts of “**assimilation**” and “**accommodation**”. Upon first contact with something new, the first response is to attempt to assign it to an existing schema. If this is not possible, a new schema is created, or, the organism adapts and further develops itself (Piaget 1970).

David Rumelhart, Donald Norman (1978) and Walter Cannon (1968) grasp the process of changing schemas with the following concepts:

- **Tuning**; a temporary adaptation or expansion of an existing schema, e.g., in recognizing a specimen of a certain animal species because of its similarity or family resemblance to other specimen of the species,
- **Accretion**; a gradual and permanent modification of existing schemas that can cause a person who spends an extended period within another culture to gradually lose the ability to differentiate between previous and newly added schemas after a certain point in time,
- **Restructuring**; a sudden and massive change in existing schemas, for example through long-term exposure to discrepant experiences, conscious reflection on experiences and their content, and/or active efforts to reorganise one's own knowledge (cf. Nishida 2005 with reference to Rumelhart and Norman 1978 and Cannon 1968).

Cultural schemas, and hence guidelines for the appropriateness of social behaviour in different situations, are, just like other schemas, created through repeated experience, observation, or discussion of similar situations, (Nishida 2005, p. 403; cf. for example Abelson 1981, Chi 1981, Fiske and Taylor 1984, Hudson and Nelson 1983, Hudson and Shapiro 1991, Mandler 1984, Minsky 1977, Schank and Abelson 1977, Taylor and Crocker 1981; Turner 1994). This not only facilitates social interaction in action contexts, such as a purchasing situation (*order, receive, pay*), and communication, it also influences problem-solving approaches or emotional and affective reactions to specific processes. Thus, among other things, cultural schemas form the **basis of a “complex cognitive structure that underlies our behaviour.”** (Nishida 2005, p. 404) and can be defined as follows:

Cultural schemas are generalized collections of knowledge that we store in our memory through experiences in our own culture. Cultural schemas contain general information about familiar situations and behavioural rules as well as information about ourselves and people around us. Cultural schemas also contain knowledge about facts we have been taught in school or strategies for problem solving, and emotional or affective experiences that are often found in our culture. (cf. *ibid.*)

Chase and Ericsson (1982) see cultural schemas as **orientation guides for information processing**; they postulate that cultural schemas are associated with the organisation of information into “meaningful chunks” (Nishida 2005, p. 409). This is for example supported by studies indicating that having prior knowledge about a certain topic enables a person to remember new information about this topic more effectively, since the information is already structured by existing schemas. Schemas can therefore be understood as a condition of possibility to perceiving situational contexts.

With regard to social appropriateness, the acquisition of *primary social interaction* schemas (see sect. 4.1.1.1) is necessary to adapt to local standards (Nishida 2005, p. 410). Cultural schemas also explain why there are culture-specific judgements of appropriateness that diverge from one another: depending on local regularities, different schemas are adopted by people, generating different expectations and expectations of expectations, and leading to irritation when these expectations are disappointed. For a more detailed discussion of each type of schema, such as fact and concept schemas, person schemas, self schemas, role schemas, procedure, strategy, emotion and context schemas—the latter of which contain specific rules of appropriateness, see also sect. 4.1.1.1.

With regard to technical systems and cultural context, artificial assistants operating in cross-cultural contexts might need to be designed to offer options to switch between different schemas, or, for self-learning systems it could be important to pay attention to the cultural standards the training data is soaked with—that is to say: what kind of schemata the system imitates.

The ability of a person to adapt their behaviour in culturally diverse interactions can be grasped by the concept of **Cultural intelligence**. According to Soon Ang and Linn van Dyne (2008) Cultural Intelligence has behavioural, motivational, and metacognitive aspects and can be quantified on a scale with the following dimensions:

1. *Drive*: A person’s interest in functioning effectively in different cultural settings. Drive is divided into intrinsic and extrinsic interest, as well as *self-efficacy* (confidence in one’s own ability).

2. *Knowledge*: A person's knowledge of cultures. This encompasses knowledge of economic and legal systems, interpersonal knowledge of values and norms in social interactions, religious attitudes.
3. *Strategy*: The way in which *meaning* is derived from an unknown situation. This encompasses awareness of one's own cultural knowledge, the ability to plan ahead, and the ability to adapt to deviations from expectations.
4. *Action*: The ability of a person to adapt appropriately to the verbal and non-verbal behaviour of another culture. This encompasses gestures, facial expressions, tone of voice, etc.

There may be functional equivalents to these dimensions that could be implemented in technical systems to make them 'culturally intelligent', especially concerning Knowledge and Action, while aspects of Drive and Strategy such as intrinsic interest and deriving meaning might be categories that could have technical equivalents but will lack the quality they have in living beings (see also chap. 7). Nonetheless these could be dimensions to take into consideration if one plans a 'culturally intelligent' system.

Social psychologist Geert Hofstede's **Cultural Dimensions Theory** which was further developed by his son Gert-Jan Hofstede and others (cf. for example Mascarenhas et al. 2013), explores the possibilities of measuring cultural differences empirically.

The authors postulate that the unwritten rules of human cultures, which Hofstede describes as "the collective programming of the mind" (Hofstede 2001, p. 1) are responsible for regulating human social behaviour. Social systems are conceived of as systems that function effectively because humans tend to act and behave according to socially acquired expectations and trust in certain probable courses of events: "I predict that [...] Mrs. X will be in the office at 8:25 A.M. tomorrow; that the taxi driver will take me to the station and not somewhere else if I ask him; that all members of the family will come if I ring the dinner bell." (Hofstede 2001, p. 2) For each prediction being made in terms of the probability of certain courses of events manifesting, the predicting subject considers both the involved people and the **situation** (ibid.). The better someone knows the "**mental program**" or "**mental programming**" of an interaction partner, the greater the probability of a prediction being correct (ibid.)—or an expectation to be justified. However, since **mental programs** are not directly observable neither for nor in humans or technical systems, Hofstede instead proposes using observable behaviour—i.e., **words and deeds**—to draw conclusions about the inaccessible

dimension of mental programming. To **make cultural characteristics measurable**, Hofstede (2015) proceeds on the assumption that although cultural values such as *'family is more important than work'* and motivations such as *'therefore I will ask for a day off to attend a funeral'* are rather abstract and hard to observe, they can partly be inferred from observables at the surface of cultural contexts, for example by asking people about their opinions and their usual practices and interpreting the results. A visible behaviour practice such as bowing when your boss enters the room (see e.g. Mascarenhas et al. 2016, p. 934) can, according to this theory also be interpreted as a symptom of certain underlying values.

Hofstede distinguishes between:

- a **universal level** of mental programs, shared by almost all or at least most people—as well as some species of animals,
- a **collective level**, shared by some but not all people; this level contains culture including language, proxemics, and ritualized types of behaviour, for example regarding food consumption or sexuality (cf. *ibid.*),
- an **individual level** that is not shared with anyone else.

And these are the **quantifiable culture dimensions** he postulates:

1. **Individualism vs. collectivism (IDV)**: How strong is in-group loyalty, how strongly is the individual identity based on belonging to a group or how strongly are individual interests guiding behaviour?
2. **Power distance index (PDI)**: How do people in lower hierarchical positions perceive and accept inequality?
3. **Masculinity vs. femininity (MAS)**: How much do gender ascribed preferences, for example between materialism and an idealisation of assertiveness and an appreciation for quality of life and cooperation differ in people performing masculine and feminine gender roles?
4. **Uncertainty avoidance index (UAI)**: How strong is the threat from unstructured situations/situations perceived as uncertain felt to be?
5. **Long-term vs. short-term orientation (LTO)**: What is the scale of anticipation concerning a time horizon? How much time is required to build relationships?
6. **Indulgence vs. Restraint (IND)**: How free are people to fulfil human desires?

Measuring these dimensions allows to represent essential differences in the social behaviour of members of different cultures, as well as to infer from that their *collective programming of the mind* or describe their **habitus** (for the similarities between Hofstede's formulation and Bourdieu's concept of habitus, see also Hofstede 2001, p. 4).

Mascarenhas et al. (2013) use parts of this framework to propose a SOCIAL IMPORTANCE DYNAMICS MODEL (SID) that has been used to create socio-cultural virtual agents (see sect. 4.2.2).

4.2.1.2.1 An example of theoretical multitude: More cultural dimensions and typologies

Besides Hofstede's quite popular cultural dimensions, other theories have been developed to study quantifiable and non-quantifiable cultural dimensions, some of which differ from Hofstede's work, some of which build on it, and some of which preceded it. Some of these approaches are briefly listed below as a repository of inspiration and as an example of how multifaceted and manifold the theories that can be considered when trying to build sociosensitive or socioactive systems are. As pointed out before, in the proposed terminology, cultural aspects are just one subgenre of the «Situational Context's» subcriterion <place>. Even in this fraction of the overall factors and criteria of social appropriateness we find a multitude of approaches to the topic, of which, for example, only using *one* aspect, such as *one* cultural dimension, from *one* of the theories mentioned below from a subfield of *one* scientific discipline can lead to the development of a sociocultural model such as the SID model as discussed in detail in sect. 4.2.2. This is to illustrate just how rich theoretical grasping of human-human interaction is, how broadly on the one hand, and how fine-grained on the other, we might need to look into it to find parts that can be transferred to the design of technical objects or of human-machine interaction in general. The following list can also be understood as a reminder to the necessity of an integrative approach to research: with a goal in mind for technological development, it can be useful to involve experts in a related field who will provide a theoretical overview.¹¹

- Based on work by Clyde Kluckhohn (Kluckhohn 1951), Florence Kluckhohn and Fred Strodtbeck (1961) propose the following dimensions of *value orientation* in action theory (cf. Hills 2002):
 1. **Time orientation of people:** Concerns orientation towards the past, present, or future.

¹¹That being said, the following list is also an example of doing just that as it is a summary of a report by Ulrich Höbner (2019) written in the context of *poliTE*.

2. **Relationship between man and nature:** Concerns the question of harmony with nature, submission, domination vs. living harmoniously.
 3. **Relationship between people and other people:** Concerns the question of social hierarchy vs. equality and hence whether a position in society can be earned through merit (as opposed to birth right, for example).
 4. **Activity orientation of people:** Concerns people's potential for growth: acceptance of an existing situation or work towards development and openness to experiences.
 5. **Essence of human nature:** Concerns the conception of whether human nature is “good”, “bad”, “mixed”, or “neutral” and whether it can be changed or not.
- Social psychologist Shalom Schwartz with his THEORY OF CULTURAL VALUES ORIENTATIONS proposes seven cultural value orientations that form three cultural value dimensions (Schwartz 2006) or *cultural-level value oppositions*:

Embeddedness vs. (intellectual/affective) autonomy

Mastery vs. harmony

Hierarchy vs. egalitarianism

In his THEORY OF BASIC VALUES Schwartz (2012) identifies ten individual values that do not necessarily reflect those of the culture. An individual may exist within a culture and reproduce its values but can also have conflicting individual values. An example of this could be someone who lives and was socialized within a culture that attaches great importance to hierarchy but individually believes that all people are equal and accordingly treats subordinates (e.g., at work) as equals or without a strong hierarchical gradient (cf. Dragolov and Boehnke 2015). For a distinction between values from attitudes, beliefs, norms, and traits and for an explanation of why there seems to be a “similar hierarchical order” of “average value priorities of most societal groups” (Schwartz 2012, p. 2) see also Schwartz (2012).

- Anthropologist Edward T. Hall's work contains up to six dimensions (cf. Rogers et al. 2002), some of which are:
 1. **Temporal orientation** (Hall 1959): Hall distinguishes between *monochronic* and *polychronic* conceptions of time. According to a monochronic conception of time (prevalent in many Western cultures), time is understood as a resource that is consumed. A polychronic conception of time views

time as unplannable. Activities are interrupted when something that is considered more relevant happens, then resumed later.

2. **Spatial orientation** (Hall 1966): Hall's spatial orientation, sometimes described as *proxemics*, examines the topic of physical distance between interacting persons in different social contexts such as intimate, personal, social, and public interaction.
 3. **Contextual orientation** (Hall 1976): The dimension of contextual orientation includes the idea of *extension transference* and can be used to describe interpersonal communication between individuals. *Low-context cultures* are characterized by the fact that little or no contextual information is needed to understand a message, and all relevant information is represented verbally (or in writing). In *high-context cultures* a large amount of information is conveyed through context, for example through *prosody* or *gestures*, but also through the communication situation itself. For example, messages containing criticism that may endanger relationships may not be communicated directly.
- Trompenaars and Hampden-Turner (1993) propose, the following dimensions, among others (such as Individualism—Communitarianism):
 1. **Universalism versus particularism**: Universalism is characterized by a rigid interpretation of rules, laws, and norms. There are no exceptions, laws always apply to everyone. Particularism represents the opposite stance, where exceptions are permissible (e.g., for celebrities, etc.). This dimension is comparable to Hofstede's dimension of power distance in some regards.
 2. **Neutral versus affective cultures**: Here, neutrality means neutrality in terms of emotion and its expression. A neutral culture does not include emotional discussions, emotions are not publicly displayed, and expressions of emotion (e.g., through gestures or speech volume) are avoided. By contrast, in an emotional society, expressing and acting out emotions is also appropriate in public. There are parallels with Hall's dimension of spatial orientation (1966).
 3. **Specificity versus diffuseness**: Specificity in a society refers to the separation of different areas of life. Work and private life are separate, and privacy is granted special protection. Diffuseness describes the mixing of different areas of life. Nevertheless, roles remain relatively constant throughout (e.g., it is appropriate to invite your boss to a party, but he or she is still your boss there). This dimension also reflects Hall's spatial orientation.
 4. **Achievement versus ascription**: This dimension focuses on the accessibility of status, recognition, and reputation in performance-oriented

(achievement) vs. status-oriented (ascription) cultures. In a performance-oriented culture, social advancement can be achieved through work and effort. In a status-oriented culture, more uncontrollable factors such as gender or caste play a determining role.

5. **Sequential time versus synchronous time:** Hall's monochronic cultures are comparable to the cultures described here as having a sequential conception of time; they are characterized by linear processing of tasks. In contrast to these, Hall defines polychronic cultures, which here are described as having a synchronous conception of time. In these cultures, tasks are carried out in parallel.
 6. **Internal versus external control:** Belief in internal control describes the conception that people can control their environment—natural, technical, and social. External control refers to the conception that there are uncontrollable (higher) powers that influence events, and people are subject to a “destiny” to which they must submit. Here, parallels can be drawn to the dimension of “relationship with nature” mentioned above.
- The GLOBE study was performed as part of the *Global Leadership and Organisational Behaviour Effectiveness (GLOBE) Research Program*, founded by Robert J. House. The results (House et al. 2004) coincide with Hofstede's cultural dimensions in some regards:
 1. **Performance orientation:** Are dedication and the pursuit of improvement rewarded by society?
 2. **Assertiveness:** Is competitive and confrontational behaviour rewarded by society?
 3. **Humane orientation:** Are fairness, altruism, generosity, etc. rewarded by society?
 4. **Gender egalitarianism:** Does society seek to achieve gender equality? (based on Hofstede's dimension of masculinity/femininity)
 5. **Future orientation:** Is future-oriented behaviour (e.g., through long-term planning) rewarded by society?
 6. **Organisational individualism vs. collectivism:** Are (institutional) collaboration and collective resource distribution rewarded? Are networks formed to conduct exchanges?
 7. **Societal individualism vs. collectivism:** Do individuals define themselves through their belonging to a (social) group and are they loyal to this group? This dimension reflects a differentiation into ingroups and outgroups.
 8. **Power distance:** Similar to Hofstede (see above).
 9. **Uncertainty avoidance:** Similar to Hofstede (see above).

- Culture, communication and language theorist Richard D. Lewis (1996) in his **LEWIS MODEL OF CROSS-CULTURAL COMMUNICATION** presents fundamental types of cultures; hybrid combinational forms are possible. Regarding social appropriateness these differences can be highly relevant.
 1. Persons in **linear-active cultures** (e.g., British, US American, German, or Swiss cultures) are strongly task-oriented and organised, diligent in their work, with strict time management and planning needs.
 2. Persons in **multi-active cultures** (e.g., Romance, Latin American, and Arabic cultures) are strongly person oriented. Interpersonal relationships are seen to be important for collaboration and are given priority over strict time constraints and the performance of tasks. Work is characterized by flexible and simultaneous task performance; communication styles are expressive and lively.
 3. Persons in **reactive cultures** (e.g., Finnish, Scandinavian, and many Asian cultures) are reserved, polite, and modest. Communication styles are characterized by active and concentrated listening, harmony, consideration of needs, and the avoidance of confrontation (including saving face of all participants, see also sect. 4.4.1.6) are granted high importance.
- Alexander Thomas (2016) designed a series of cultural standards within the framework of qualitative binational studies comparing different cultures. Seven German cultural standards from the perspective of many foreign partners are listed below as an example (cf. Schroll-Machl 2016); these standards should be viewed as tendencies rather than established facts:
 1. **Task orientation:** People and facts are separated. Especially in professional life, person interests are set aside, and emotional issues are disregarded. This makes work and communication styles efficient and target-oriented but can also represent an obstacle for collaboration if fundamental problems at a relationship level are ignored.
 2. **Valuing of structures and rules:** Rules are strictly always followed by everyone, and planning systems are necessarily considered. If somebody violates a rule, it is socially appropriate for this to be sanctioned by uninvolved observers (e.g., rebukes/complaints if somebody runs a red light). Special treatment is only granted informally and under certain conditions. Reduced flexibility can lead to problems if there are unanticipated events.
 3. **Rule-oriented, internalized control:** People socialized within the German culture are characterized by a strongly rule-oriented manner, identify strongly with concepts, activities, etc., that are perceived to be good and appropriate, and consistently implement them based on internal motivation for their perceived moral value even in the face of (internal) resistance.

The consequences include high reliability and self-discipline, but also low flexibility.

4. **Time planning:** German culture has a monochronic conception of time (see above): time is viewed as a finite resource that must be carefully used and precisely planned (including for personal matters). Deviations from a schedule are only accepted for good reason (and if necessary, with an apology), untimeliness is socially appropriate, and interruptions in work on a task are undesirable.
5. **Separation of different domains of personality and life.** Work and private life are strictly separated in German culture, and work is not expected to be interrupted by personal matters. A social relationship is not required for collaboration, spending leisure time with colleagues is optional. Different social groups (e.g., circles of friends) therefore tend to remain separate.
6. **Low-context communication style:** In German culture, people communicate directly and clearly, including about problems. Politeness plays a secondary role to clarity of information (compare with low-context cultures, see above). This makes communication styles extremely efficient and effective but can seem unfriendly and inappropriate to more strongly relationship-oriented cultures.
7. **Individualism:** Self-reliance and independence are often valued more highly in German culture than peer groups, personal goals are more important than collective interests, decisions are made by individuals themselves. German culture is therefore described as an individualistic culture. One disadvantage is that support must be actively sought to address personal problems, including assistance of a professional nature (Schroll-Machl 2016).

4.2.2 «Situational Context» (culture) and sociosensitive/ socioactive artificial assistants

Technical systems can be seen as performing several roles regarding Situational Context, for example:

- (a) Technical systems might be seen as requisites that help determine how a situation is defined by a human and may be designed accordingly. For example, if the goal is to influence or alter situation perception a system can help in doing that—be it in health care contexts in which a person's situation evaluation is desired to be changed, in a context in which situation evaluation is mediated technically, for example where danger is signalled by a technical medium, or

- any other scenarios. Questions leading the development process could be: does the assistant somehow influence interacting agents' situation perception? Is it supposed to be?
- (b) Technical systems might be seen as participating a predefined situation. Questions could be: is the assistant designed to fit the situational context? Is it supposed to be?
 - (c) Technical systems might be seen as observing a situation. Questions could be: What does the assistant need to understand the situation? What can be possible misconceptions? Who is to decide what a misconception is and where, when and by whom is data being interpreted? For example, in machine learning, an unsupervised learning system is not so much designed to produce something "wrong" or "right" as a supervised system, a pattern-detecting system might observe a pattern unfamiliar to a human observer and can come to another conclusion concerning the situation.
 - (d) Technical systems might simulate a situation to help a person orient themselves in such a situation should it manifest in their life. Examples are technically mediated trainings such as in-flight simulation or intercultural training.

The following example of the SOCIAL IMPORTANCE DYNAMICS MODEL illustrates the use of selected theorems taken from theories associable with FASA MODEL factors and factor criteria in sociosensitive and socioactive systems development. The SID model has been used in training contexts such as mentioned in (d), and it offers an architecture to design agents in a way so they can quantify the social appropriateness of certain action options a system has the possibility to choose between.

With a model based on a theory of SOCIAL IMPORTANCE DYNAMICS, building upon some of Hofstede's work and Theodor Kemper's STATUS-POWER THEORY (see sect. 4.4.1.5), Mascarenhas et al. (2016) propose, it should be possible to create virtual agents that are 'aware' of the social consequences of their actions, and which can take certain aspects of a culture into account and choose their courses of actions accordingly (Mascarenhas et al. 2016, p. 933). This is attempted by simulating human social behaviour according to the dimension of individualism vs. collectivism with the SID model. The authors remark that previous virtual agents have mainly focused on taking **language** (Johnson et al. 2005), **dialect** (Finkelstein et al. 2013), **pauses and overlaps in speech** (Endrass et al. 2009), **posture** (Endrass et al. 2011), and **proxemics and eye gaze** (Jan et al. 2007) into account (Mascarenhas et al. 2016, p. 932). Instead of focussing on those features, the authors promise to focus on "cognitive biases in judging the appropriateness of social actions in everyday social interaction" (Mascarenhas et al. 2016, p. 932).

Mascarenhas et al. identify three levels that are important for the development of sociocultural agents and regarding situations: an **interaction** level, a **group** level, and a **society** level (Mascarenhas et al. 2016, 937ff.).

1. **Interaction:** At the interaction level, the **ritual** (cf. also Wulf 2022) must be taken into account, understood as an at least dyadic interaction process involving shared attention, possibly with symbolic elements, together with the physical context (place) and the participants of this ritual; the physical context and properties of the participants also influence the ritual itself (Mascarenhas et al. 2016, p. 937 with reference to Degens et al. 2014); here, ritual ultimately refers to every possible social (and possibly symbolic) action, which is partly covered by the concept of situation in the FASA MODEL.
2. **Group:** At the group level, “**moral circles**” (Wilson 2007) need to be considered, which determine to whom certain rights and obligations are attributed, for example, and which contain three further elements: a) members of the *moral circle*, b) their shared perception of moral attributes (relational variables), and c) rules of behaviour for the members of each *moral circle* (social norms). In the computer-based model, the members of a circle are embodied by agents with the same configuration. Relational variables describe aspects of a relationship such as prestige or reputation, and the sum of all relational variables with positive connotations constitutes the ‘social importance’ of an interaction participant. Social norms determine which actions should and may be performed in certain contexts and which are prohibited (Hollander and Wu 2011). In every moral community, there are different and overlapping norms for behaviour that govern its appropriateness. (Note that the notion of “moral circles” corresponds to the factor Standards of Customary Practice)
3. **Society:** At the society level, the behaviour of the society’s members is primarily regulated by cultural **meta-norms**. The cultural background might for example determine which moral communities and which social norms are viewed as more or less important, for example: is it more important to accept a party invitation from your boss or from your friends?

Fig. 4.4 (reproduced from Mascarenhas et al. 2016, p. 939) shows how the SID model is integrated within a belief-desire-intentions (BDI) **architecture** (see e.g. Rao and Georff 1995) to improve the social intelligence of the BDI agent.

Beliefs stand for knowledge about the environment, the self, and other agents, and are represented by logical predicates. *Desires* denote the agent’s motivations and formalize the world state that the agent wants to make happen. In this model, a goal G has the values (1) name, (2) preconditions, (3) success conditions, (4) importance of success. *Intentions* represent goals pursued by the agent at the

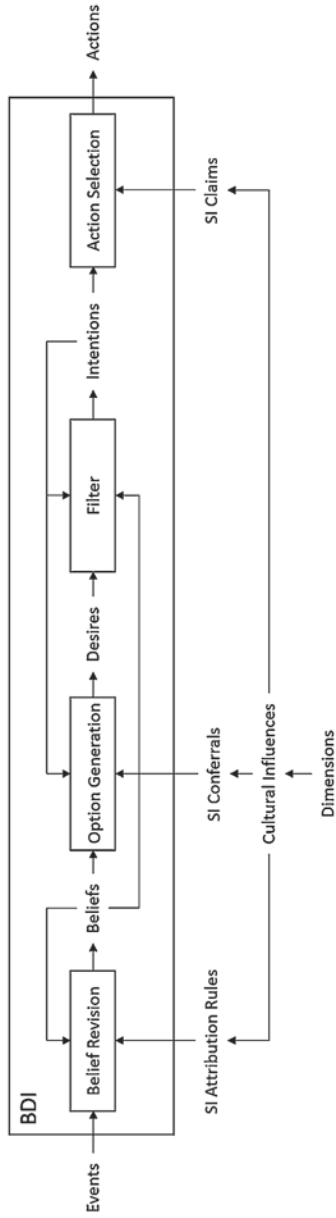


Fig. 4.4 Social importance model applied to general BDI architecture (Mascarenhas et al. 2016, p. 939)

current point in time, consisting of a particular goal and an action plan to achieve its success conditions.

The authors now insert the relation of *social importance* (SI) into this BDI algorithm. Its value influences whether an action should be carried out or dropped by the agent. The agent starts with its own beliefs (*Bself*) and the beliefs of its partner (*Bag*) (Mascarenhas et al. 2016, p. 940). Initially, the agent's own beliefs are assumed to be also existent in its partner; then the social importance, the agent's assumptions about the beliefs of its partner, and its own beliefs are revised based on characteristics observed by the agent. The SI of an interaction partner is represented by the formula $SI([x])=[y]$, where x denotes the target agent's name and y is the SI value attributed to it using so-called *SI attribution rules* that encode "shared cultural knowledge of how different factors influence the SI of other individuals" (ibid. p. 941). In the simple example presented here, the factors "is a person" and "belongs to ingroup or outgroup" are quantified to calculate the SI (ibid.). The SI attribution rules also contain a mechanism that deducts a certain value for inappropriate behaviour. The model does not consider the agent's desire to collect SI by displaying status-enhancing behaviour; however, this is possible in principle according to the authors (ibid. p. 942f.). The goal of the model as presented here is to code socially appropriate behaviour. The SI attribution to interaction partners occurs via:

- A set of preconditions that determine in which contexts SI attribution is appropriate (C)
- A name for the action that symbolizes the status conferral / SI conferral (A)
- A target agent to whom the conferral is made (T)
- A conferred social importance value (V)

Actions such as *giving advice*, *giving directions*, *offering help*, *greeting*, *showing gratitude*, and *applauding* are associated with an SI conferral. As shown in example 5, certain demands or requests are assigned values (V) to enable the agent to calculate whether its attributed SI in a given context is sufficiently high to make a request, which is also assigned a value.

EXAMPLE 5: USAGE OF THE SI MODEL

Consider the following **example**: John and Paul (InGroup) are in a bar together with User (OutGroup). User and Paul are looking for a place to sleep.

The possible actions are: `askDirectionToHotel(V = 10)` and `askToSleepOver (V = 60)`.

Paul now calculates whether $SI(Paul) \geq 60$ is true in *BJohn*, i.e., with regard to his model on John's mental state, Paul checks to see whether John attributes enough SI to him to possibly accept the request to sleep over at his house (Mascarenhas et al. 2016, p. 945). In the example, within his model of John's mental state *BJohn*, Paul expects John to attribute $SI(Paul) = 25$ to him, based on $Person = 20$; $InGroup = 5$. Since this is not enough to request an overnight stay, because $25 < 60$, he instead asks John for directions to a hotel.

In simplified terms, a cultural difference can be represented in this model by assigning different values to the attribute *InGroup* or *OutGroup* to reflect different values along the culture dimension of individualism (IDV) versus collectivism (COL). In a more collectivistically organised culture, agents expect a stronger evaluation of the difference between ingroup and outgroup, i.e., different behaviours can be expected depending on whether the agent is interacting with an ingroup or outgroup member. In a more individualistic culture, it can instead be expected that the interacting agents will be treated more equally regardless of their belonging to a particular group (Mascarenhas et al. 2016, 945f.). In the model, this would result in different values when evaluating the beliefs and mental states of an interaction partner. It would also lead to evaluation differences regarding potential actions to be performed: greeting, addressing, or joining an unfamiliar group would for example tend to be seen as more appropriate in individualistically organised cultures than in collectivistically organised cultures (Mascarenhas et al. 2016, 946f.). Conversely, the performance of such acts, being associated with a lower cost, would be less socially worthwhile and significant. The authors use this to explain why the greeting "How are you?" in the strongly individualistic American culture is often misunderstood as genuine concern by interaction partners socialized within other cultures and in some cases met with a detailed response (ibid., 946f.).

An example of a more complex architecture, where the SI is implemented into "FAtiMA [14], a well-established architecture for emotionally intelligent agents that follows the OCC Appraisal Theory [15]" (Degens et al. 2013, p. 5) is to be found when looking at the intercultural training application called TRAVELLER, aimed at a target group of 18–25-year-olds, in which *critical incidents* of an interaction with virtual agents can be replayed (Degens et al. 2013). The goal of the application is to demonstrate cultural differences by assigning different SI values to players based on their actions, causing the virtual agents to behave—and judge (raising eyebrows, monosyllabic responses)—in different ways accordingly. The

program was tested in Portugal, the supposedly most collectivistic European country, and in Holland, the supposedly most individualistic European country.

Other culture-aware agents, such as ELECT BILAT (*Bilateral Negotiation Trainer*) (Hill et al. 2006), are for example used in training for soldiers with the goal of teaching language dialects, appropriate gestures, and negotiation tactics (Hill et al. 2006; Johnson et al. 2005). For negotiations, PURB (Haim et al. 2012) was developed based on COLORED TRAILS (Gal et al. 2011; Grosz et al. 2004) as an agent that observes, processes and negotiates human social behaviour during negotiation talks; here, a probabilistic model of human action was used. The CAB MODEL (Solomon et al. 2008) which models culturally affected behaviour and THESPIAN (Si et al. 2006) which models socially normative behaviour were, just as the TRAVELLER, equipped with a *theory of mind mechanism* based on the PSYCH-SIM FRAMEWORK (see Mascarenhas et al. 2016, p. 935). Degens et al. (2014) developed a conceptual model for sociocultural agents based on Hofstede's dimensions of *individualism vs. collectivism* and *power distance*. ORIENT (Aylett et al. 2009) is designed as a game to teach intercultural empathy to teenagers on the basis of a model of social behaviour whose architecture is partly based on Hofstede's work as well. Some other models (Insafutdinov et al. 2016; Torta et al. 2013) are sociosensitive or culturally sensitive to the extent that they take proxemics into account. Nick Degens et al. (2017) offer further suggestions based on the SID model on how culturally coded aspects of behaviour could be integrated into virtual characters. They performed various experiments showing that test subjects with different cultural backgrounds evaluate the social behaviour of a virtual character differently. The results are intended to help to design technology that reflects how meaning is attributed differently by different local cultures.

4.3 The «Individual Specifics» factor

«Individual Specifics»¹² refers to individual influences on the construction and/or perception of social appropriateness in interpersonal interactions, for example

- the overall physical, psychological, and cognitive constitution of interacting agents, as well as their actual state of being,
- the personality structure of the interacting agents,
- their age, gender, individual social roles and positions in society, individual lifestyles (fashion, food, etc.).

¹² In the terminology of differential psychology, this factor encompasses inter- and intraindividual differences (Stemmler et al. 2016).

One example concerning the actual state of a person in a certain situation could be their current emotional state or their being influenced by outside stimuli. It has for example been demonstrated empirically that people who are exposed to an odorous substance (chemically synthesized hedione) sanction their interaction partner's social behaviour of more than people who were not exposed to the substance in game-theoretically constructed experimental situations (Berger et al. 2017; this is also relevant to the factor «Relations between Interacting Agents»). More generally and not situation bound but personality bound, specific character traits can for example be used to conclude which type of interpersonal communicative interaction style (more or less dominant, more or less hierarchically structured) certain people prefer individually, (Pratto und Sidanius 2012); cf. sect. 4.3.1.2. Both these examples show how an individual's judgement regarding the appropriateness of another's action and/or behaviour can vary according to the judge's situational or personal states of being. Additionally, appropriateness assessments can, for example, vary according to gender, age, skin colour and other personal features (Curtin 1985; Mills 2003; Tolmach Lakoff and Bucholtz 2004). Both the evaluating and the evaluated person's characteristics can influence what will be perceived as appropriate behaviour. On a meta level, it is possible to assess the appropriateness assessment itself from an ethical point of view: Do we think it is appropriate to make this or that judgment about a person's behaviour if we count in all the above-mentioned individual specifics of interacting agents?

These examples show that:

Overall, **what is perceived as appropriate for and by whom** varies from individual to individual.¹³

The relevance of individual features for questions of social appropriateness is theoretically grasped in Pierre Bourdieu's comments on (individual) habitus and (individual) lifestyles (Bourdieu 1982), as well as in the fundamental conviction of theories of SOCIAL SIGNAL PROCESSING, namely that the human ability to behave socially is strongly based on (individual) processing of social signals by others (Pentland 2007; Salem and Dautenhahn 2017; Vinciarelli et al. 2009; Vinciarelli et al. 2012; Vinciarelli et al. 2017).

¹³However, variability does not imply arbitrariness, as demonstrated by the factor criteria listed below—and possibly others.

Possible factor criteria:

- **«Personal evaluation structures»:** What specific preferences do the interacting agents have?
- **«Personal (persistent and temporary) characteristics»:** What specific dispositions and/or (possibly short-term) characteristics do the interacting agents bring with them?
- **«Individual shaping of social roles»:** What specific social roles are the interacting agents ascribed to and how do they interpret them?

Examples of questions and aspects that can be assigned to each factor criterion to support an analysis of the potential social (in)appropriateness of interaction situations and guide these situations are given in Table 4.3.

Possible observables:

- Expressions of preference; judgements of the behaviour of others; spatial distance from interlocutors; content and duration of a conversation; touching when greeting, during conversation, when saying goodbye, etc.
- Specific behavioural reactions in specific situations: are interactions initiated with strangers, and if so, how? Additionally, strong behavioural reactions like trembling, sweating, etc.; permanent behaviour patterns, etc.
- Voting behaviour; party donations; eye contact; speech volume; professional attire; accessories; clothing style; diet, etc.

4.3.1 Theoretical perspectives on «Individual Specifics»

The factor «Individual Specifics» is central to all human action and behaviour and is the subject of various research and entire scientific fields, such as anthropology, (social) psychology, and—in connection with the individual configuration of social roles—sociology and various cultural sciences. In the context of social appropriateness, ‘individual specifics’ are often implicitly relevant due to this central role but are not considered as a topic in their own right. The following examples of theoretical lines of questioning show that this factor can nonetheless be fruitful for the examination of socially appropriate action and behaviour, and how.

Table 4.3 Individual Specifics. (Source: own representation)

«Individual Specifics»	
«Personal evaluation structures» What specific preference hierarchies do the interacting agents have?	Preferences, prejudices, tastes relating to social appropriateness, such as: what distance does someone personally consider appropriate during a conversation? When and how is it felt to be appropriate to talk about private matters? What degree of physical contact is considered appropriate?
	Are personal evaluation structures affected by a 'negativity bias'? Are there any other 'perception biases' or possibly pathologies?
	Are there stereotypical evaluation structures for questions of social appropriateness, and if so, which? For example: are people read as women generally perceived or expected to be more polite?
	What attitudes are there towards social positions or political, cultural, gender-related questions, etc.?
	What self-attributions are there, for example in terms of skills (e.g., self-efficacy assessment)? Also directly relating to social appropriateness, e.g.: how polite, tactful, etc. do the interacting agents consider themselves to be?
«Personal characteristics» What specific dispositions and/or (possibly short-term) characteristics do the interacting agents bring with them?	What personality traits do the interacting agents have? For example: open-hearted, shy, stubborn, cooperative, etc. but also, e.g.: how strong is the ability to tolerate ambiguity (this 'ambiguity tolerance' has on several occasions been described as an especially important personality trait of the present age as so called 'late modernity' (Reckwitz 2019). Are there any psychosocial defence mechanisms in place, and if so, which?
	What are the physical and emotional states of the interacting agents? Is a person tired, drunk, sad, pregnant? Is their judgement influenced by certain permanent or temporary special conditions?
	How pronounced is the emotional and social 'intelligence' of the interacting agents?
«Individual shaping of social roles» What specific social roles do the interacting agents have and how do they shape them?	What class, what milieu, etc. do the interacting agents belong to?
	What professions and hobbies do the interacting agents have?
	As what gender are the interacting agents usually read?
	Which physical or mental capabilities do agents have?
	What ages are the interacting agents? Are they for example read as 'children', 'adolescents', 'young adults', 'senior citizens', etc.?

4.3.1.1 Social cognitive theory

According to Albert Bandura (Bandura 1986, 1989, 1997, 2006), for example, socially appropriate behaviour is learned individually and socially. His SOCIAL COGNITIVE THEORY (SCT) emphasizes the individual agency of actors—in the sense of **self-efficacy**, a gradually and deliberately adjustable conviction of being able to successfully perform an action. This marks a departure from theories of learning that reduce learning processes to merely environmental dependence or internal drive. SCT proposes a reciprocal, triadic relationship between environmental influences, cognitive, emotional, and biological processes and the behaviour displayed by an individual. According to this theory, learning processes that involve observing the behaviour of human role models and later reproducing it in situations judged to be suitable play a central role in the acquisition of socially appropriate behaviour—this process is also described as **social learning** (Bandura 2015). Both the cognitive, emotional, and biological processes mentioned above and this reproduction behaviour in the context of social learning are related to the factor «Individual Specifics» in the context of social appropriateness: like in any other learning process, what is learned as socially appropriate behaviour and how depends at the very least on individual specificities of the learner, e.g., attention or memory skills, fine motor skills, balance, emotional constitution, etc.

4.3.1.2 Social dominance theory

In addition to research into learning processes, (social) psychology research into social behaviour and personality traits can be cited to illustrate the «Individual Specifics» factor (the mentioned research is also relevant to «Relations between Interacting Agents»). So-called *social dominance orientation* (SDO), which is part of SOCIAL DOMINANCE THEORY (SDT), can be used to make significant predictions about the extent to which group members will demonstrate particular social attitudes and social behaviour (Ho et al. 2012; Pratto et al. 1994; Pratto et al. 2016; Pratto and Sidanius 2012). It consists of two factors: a **dominance factor** (SDO-D: how readily do group members dominate other groups?) and an **equality factor** (SDO-E: how much are ingroup inequalities preferred?). The SDO correlates with other **personality traits**, such as the traits queried in the psychological OCEAN MODEL (also known as *The Big Five personality traits*). Regarding social appropriateness, the predictive power of the SDO scale is especially clear in judgements of social behaviour of so-called ‘minority groups’: for example, the behaviour of homosexual persons is generally rated as more **inappropriate** by people with a high SDO than the same behaviour by heterosexual persons. The SDO questionnaire also captures characteristics that can be important to consider

for socially appropriate behaviour, such as the desire for equality or a power imbalance regarding both ingroups and outgroups.

4.3.1.3 On the logic of appropriateness as a Decision-Making theory

A decision theory approach by Mark Weber, Shirli Kopelman, and David Messick is especially illustrative of the «Individual Specifics» factor in social appropriateness. Here, decision-making is examined in so-called ‘social dilemmas’, which are understood as situations where individually reasonable behaviour produces results that are worse for all participants than the results that could be achieved with individually less reasonable behaviour—individual rationality leads to collective irrationality (Weber et al. 2004, p. 281). Building on March’s thoughts regarding a **logic of appropriateness** (1994), Weber, Kopelman, and Messick propose that decisions are formed by the **recognition of the situation** («Situational Context»), the **identity of the agent** («Individual Specifics»), and the application of **specific rules of action** («Standards of Customary Practice»), in other words—in contrast to radical *rational choice* approaches—decisions are not only oriented by the goal of achieving the greatest benefit. In this sense, an action is appropriate if it is an element of a set of possible actions formed by the factors of situation recognition and identity, and if it is performed in accordance with certain specific rules of action for such a situation. Therefore, besides the «Individual Specifics» factor, this theory also points to influence from the «Situational Context» and «Standards of Customary Practice» factors.

However, Weber, Kopelman, and Messick focus more specifically on the factors on which the **subjective definition of a decision-making situation** depends, since these factors essentially pre-structure the set of possible actions and thus the **space of rules** that can be applied to a potential decision-making situation, and hence also the **space of appropriate actions**.

The definition of the situation informs the person about the norms, expectations, rules, learned behaviours, skills, and possible strategies that are relevant. (...) The definition of the situation suggests a choice set. Choosing among the options, we contend, is a rule-directed exercise. (ibid. p. 285)

In the sense, the factors «Individual Specifics» and «Situational Context» can be considered primary compared to the factor «Standards of Customary Practice» in this discussion. As the cited passage also makes clear, the selection from the set of possible actions is also understood as a rule-based practice. Weber, Kopelman,

and Messick do not discuss any further which specific rules might apply as social norms in potential decision-making situations, but they differentiate this rule-based practice: it is divided into two separate practices, a **shallow** rule orientation in the selection of possible actions (i.e., implicit rule orientation or rule of thumb) and a **fine** rule orientation, where the selection of the rule itself is also the subject of conscious deliberation. The following diagram (cf. Figure 4.5) summarizes the decision-making factors in social dilemmas and their relationships to one another (ibid. p. 284).

Finally, an analysis of the literature on decisions in social dilemmas (cf. Figure 4.6) provides Weber, Kopelman, and Messick with variables that offer more precise insight into the **decisive factors of the subjective definition of the situation** (ibid. p. 287):

According to this diagram and the above discussion, the **subjective definition of a decision-making situation** and therefore the set of possible actions and the space of appropriate actions depends on the actor's **identity** and **recognition of the situation**. The **actor's identity** is understood as their specific **individuality**, which is constituted by personality traits, gender, personal history, and individual social motives, and therefore encompasses several key factor criteria of the «Individual Specifics» factor. The actor's **recognition** of the situation, on the other hand, is tied to specific **situational factors**. However, to the extent that these factors must be **recognized by the actor**, they do not only depend on *perceptual factors*, i.e., the *frames* structuring the actors' perceptions and the real perceptible cause-effect relationships. This is also why we can describe the considerations of

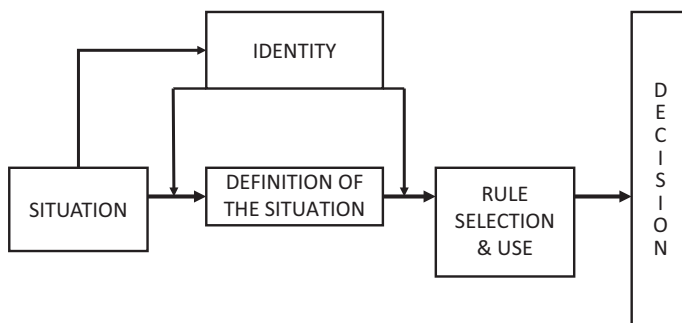


Fig. 4.5 Schematic representation of decision factors (Weber et al. 2004, p. 284)

DECISION MAKING IN SOCIAL DILEMMAS

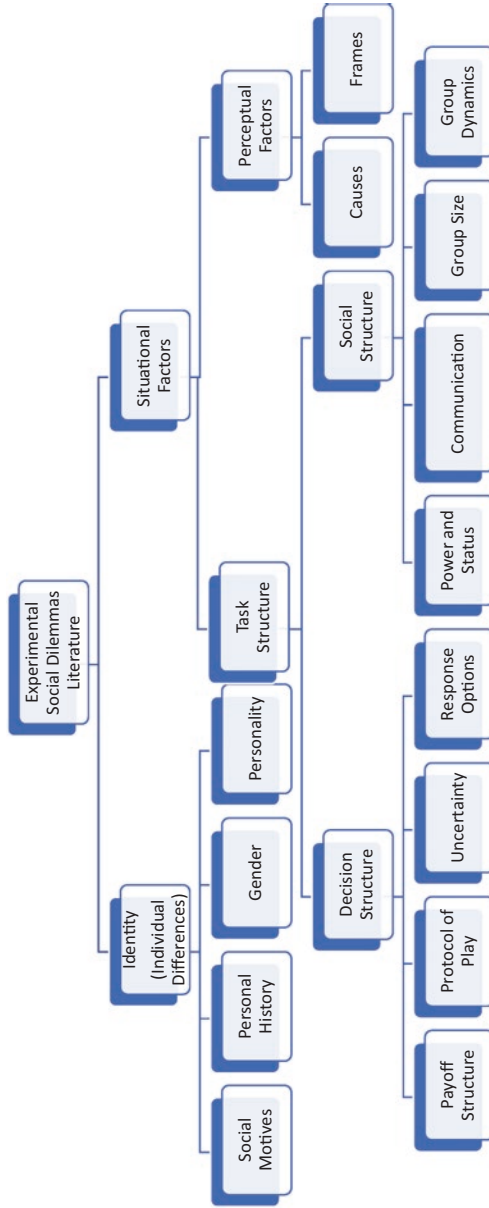


Fig. 4.6 Classification of literature analysis on social dilemmas. (Modified after Weber et al. 2004, p. 287)

Weber, Kopelmann, and Messick as methodological individualism in the sense of a methodological primacy of the individual—meaning that the individual occupies a central theoretical position—and if nothing else why they are particularly suitable for illustrating the factor «Individual Specifics». In the context of situation recognition, what must be recognized is first and foremost the object (broadly, what must be decided) of the decision (*task structure*). What must be decided is in turn determined by the essential structure of the decision object (*decision structure*), such as the benefits that can be gained from each action that could be chosen, or how much uncertainty the action is subject to, as well as the *social structure* of the decision object, such as the power and status relations involved.

4.3.2 «Individual Specifics» and sociosensitive/ socioactive artificial assistants

For artificial assistants, the factor «Individual Specifics» is crucial in the context of socially appropriate human-machine interactions. At a fundamental level, there is an **‘individualization trend’** in assistive or at least interactive technical systems—from applications to robots. Technical systems are increasingly expected to (have the ability to) adapt to the wishes and needs of singular users. Under the buzzword of ‘personalization’ (etwa Pariser 2017), this has almost become a kind of design credo and a distinctive trend within a world that is being algorithmized by comprehensive IT systems (Wiegerling 2016). Artificial assistants that can interact with users in a socially appropriate manner will need to keep step with this trend for this reason alone.

And even disregarding any trends relating to personalization, it is crucial to consider «Individual Specifics» in the design of socially appropriate human-machine interactions: as the above discussion demonstrates, social appropriateness in interpersonal interactions also heavily depends on the specific design of individual factor criteria. This also applies to human-machine interactions—even if the details may differ somewhat: for example, depending on the personality profile of the driver, an artificial driving assistant (with the end goal of ‘autonomous’ driving) might be perceived as intrusive or even disempowering (cf. e.g., current research within the framework of the project *MoFFA—Holistic model for describing the distribution and handover of tasks between human drivers and artificial driving assistants in automated and networked driving* (TU Dortmund 2020a; especially "Identifikation von Fahrertypen im Kontext des automatisierten

Fahrens TU Dortmund 2020b)¹⁴. Individual attitudes also heavily determine whether the behaviour of a robot that simulates emotional behaviour (such as the ROMAN robot (Hirth et al. 2011)) is perceived to be socially appropriate or inappropriate. In general, specific and **individual preference hierarchies** thus play a key role in the context of human-machine interactions designed to be socially appropriate. Furthermore, «Individual Specifics», understood as **personal characteristics**, have of course long played a decisive role in social robotics. Entire fields of application have been defined to reflect certain characteristics, for example in autism research (Tennyson et al. 2016) or care robots (Di Napoli and Rossi 2019). In the age of a pluralization of life forms, the demand for **individually shaped social roles** should not be underestimated, including in artificial assistants. When designing an artificial assistant, it is important not to fall behind social realities regarding role conceptions (household robots that typically only reproduce female characteristics (voice, appearance, etc.) are—aside from political and ethical objections—**badly designed products**), but certain design components also need to be taken into consideration: from the perspective of appropriateness, the way in which a *robot companion* should interact with its user depends (at the very least) on individual conceptions of specific users regarding interaction roles and cannot be deduced outright from social role models at the societal level for every potential user.

4.4 The «Relations between Interacting Agents» factor

In an interaction, the interacting agents do not meet as solitary agents but as actors who are engaged in an ongoing relationship with one another; in other words, there are «Relations between Interacting Agents». Although every individual has their own habitus and lifestyle, these aspects are also determined by social position and roles in relations to others, in general and in concrete situations. Discussions in classical *POLITENESS THEORY* by anthropological linguist Penelope Brown and social scientist Stephen Levinson for example point to social proximity/distance, the power relationship between the interacting agents, and the specific culturally coded degree of threat of speech actions as being relevant to the question of social appropriateness in interpersonal interactions (Brown and

¹⁴Own translation of project name. In the original it is ‚Holistisches Modell zur Beschreibung der Aufgabenverteilung und der Aufgabenübergabe zwischen menschlichem Fahrer und Fahrerassistenzsystemen beim automatisierten und vernetzten Fahren‘.

Levinson 1978, 2011). The relations between the interacting agents can also be observed in kinematics ('movement theory'), for example. Here, it becomes clear that the social conditions of an interaction situation even influence the behaviour of the interacting agents at the level of motor skills (Krishnan-Barman et al. 2017): for example, the perceived attractiveness or assumed group membership of the interaction partners affects the execution of simple gripping and pulling movements by an actor. A very concrete example of how the way interaction partners relate to one another is reflected institutionally, for example, is the German legal system's elaborate list of persons who are granted the right to refuse to testify in court due to their relationship with the accused.

Possible factor criteria:

- **«Familiarity or relationship aspects»:** How are the interacting agents related to one another?
- **«Intention»:** What are the interests underlying the interaction?
- **«Context»:** 'As what' is the interaction taking place?

Examples of questions and aspects that can be assigned to each factor criterion to support an analysis of the potential social (in)appropriateness of interaction situations and guide these situations are given in Table 4.4.

Possible observables:

- Physical distance during the interaction, type of greeting actions performed, form and content of the conversation, facial expressions, gestures, phenotypic features, clothing, emblems, any status symbols, etc.
- Type of action; features of the location and environment such as buildings, concrete room, including any props, etc.
- Typical frame markers, e.g., for a theatrical play: stage, audience, curtain, lighting, tickets, programme announcing the performance, etc.
- Example: for possible observables concerning «familiarity» see Pierre Feyereisen 1994.

4.4.1 Theoretical perspectives on «Relations between Interacting Agents»

The following theories and concepts are illustrative in understanding the factor «Relations between Interacting Agents».

Table 4.4 Relations between Interacting Agents. (Source: own representation)

«Relations between Interacting Agents»	
«Familiarity and relationship aspects» How are the interacting agents related to one another?	How often do they interact? How often have they already interacted? How long are they interacting for and in what way?
	What 'level of friendship' is there between the interacting agents? How strongly do they feel connected (<i>tie strength</i>)? What is the prevailing level of familiarity?
	What is the degree of reciprocity? Is there an ingroup/out-group relationship between the interacting agents?
	How are respect/affection and good will relationships arranged between the interaction partners? Are the rights and duties of agents observed and their interests fulfilled? Does the interaction contain <i>face-saving</i> efforts or shaming efforts?
	Is the interaction voluntary? Does the user/interaction partner have freedom of choice?
	How are the power dynamics between the interaction partners, or how are they attributed? (e.g., differences in social hierarchy, social distance, social diversity, or heterogeneity? Institutional power dynamics? Differences in authority/access to resources? Also relevant: status, respect, and degree of familiarity)
	What degree of ' <i>social connectedness</i> ' is there? Keywords: relationship saliency, closeness, see for a conceptual model and a design example with regard to social connectedness Visser et al. 2011
	Expectations and expectations of expectations: mutually attributed freedom of action, attributed emotional variance, attributed perception, and personality structures (e.g., preferences, pathologies)
	Economic (persuasive arrangement of goods, speed of deliveries, rating systems in sales portals)
«Intention» What are the interests underlying the interaction? (closely related to «Situational Context»)	Is there a form of <i>task imposition</i> , i.e., is a task being imposed on an interaction partner?
	Is it a confrontational or a cooperative interaction? Logistic organisation? Negotiation context?
	Is the interaction taking place in a professional or private context? How openly are personal interests presented?
	What are the possible consequences? How is the relationship between the interacting agents institutionally enshrined? Observable example: judicial rules about testimonies concerning family members (in the German legal system)

(continued)

Table 4.4 (continued)

«Relations between Interacting Agents»	
<p>«Context» ‘As what’ is the interaction taking place? Also, roughly similar to the «Situational Context» factor, here under the aspect of framing</p>	<p>Assigned meanings: ‘As what’ do the interaction partners perceive the interaction? Is there a consensus/disagreement on this?</p>

4.4.1.1 Respect and recognition

The concept of respect or recognition has proven to be relevant for the factor «Relations between Interacting Agents» (as well as for «Standards of Customary Practice» and «Type of Action, Conduct, Behaviour, or Task»). **Theories of recognition** include in particular theories that emerged from German idealism (Fichte 2013 [1796]; Hegel 1980 [1807]) and were subsequently expanded (Bedorf 2010; Habermas 1968; Honneth 1994; Siep 2014; Taylor 2009; Waldenfels 1998) to explain phenomena such as normativity, individual and collective identity, and understanding of the world in terms of interpersonal action and knowledge relations. The concept of recognition possibly represents a basic prerequisite of the phenomenon of social appropriateness in interpersonal interactions. One aspect of recognition is the topic of **respect**, which can be understood as a superordinate concept to various models and a prerequisite for their implementation. Thus, even if one is perfectly aware of the norms or cultural standards that should be applied in a certain situation, choosing to respect them is a further (deliberate) step that is necessary to produce socially appropriate behaviour. **Current interdisciplinary research on respect** draws its definition of the concept of respect from the philosophical tradition of discussing and researching concepts such as **dignity, appreciation, recognition, and esteem** (Vogt 2019). The following discussions are a summary of a more detailed report (Vogt, 2019), see also Vogt 2022 for a full article derived from the 2019 report.

Extensive philosophical foundations for research on respect were established in Immanuel Kant’s ‘The Metaphysics of Morals’ (Kant 2008, pp. 18–21). In reference to Kant’s work, Stephen L. Darwall (1977) describes respect as the deliberate consideration of another person affected by the actions of the potentially respectful person, while differentiating between (*moral*) *recognition respect* (a moral obligation to which every person is entitled) and *appraisal respect* (the recognition of positively evaluated character traits or performances). Axel Honneth (Honneth 1994) differentiates the concept of recognition into recognition,

love, and appreciation, whereas Robin S. Dillon (Dillon 1992) unites the concepts of respect and care, defining affective, conative, active, and cognitive dimensions of respect. This more philosophical research on respect diverges from **psychological**, **sociological**, and **economic** research on respect but nevertheless references some aspects of it. Anthony J. Delellis (2000) names seven aspects connotatively assumed by the concept of respect: admiration, appreciation, esteem, honour, reverence, deference, and fear. He also describes twelve objects towards which respect can be directed, including nature, other people, and civilization. Based on Darwall's bifurcation into *recognition* and *appraisal respect* (Darwall 1977), Niels van Quaquebeke, Tilman Eckloff, and Daniel C. Henrich (2007) define horizontal and vertical respect. For horizontal respect, the "unconditional consideration of a person as an equal counterpart regardless of their characteristics and performances" plays a central role; "horizontal respect is categorial in two stages—respect and disrespect" (cf. Vogt 2019, p. 5; 2022, own translation). Vertical respect, on the other hand, factors in the characteristics and behaviour of the respected person. The concept of respect is in some works also understood as *employee recognition* (cf. Bruin & Dugas 2008). Furthermore, it is worth noting that parts of respect research, especially as means of developing respectful leadership techniques by consulting with interaction partners, seek to determine in what cases employees feel as they are being treated respectfully. Within these approaches respectful behaviour is understood to have only been performed if interaction partners have *felt* respected (Quaquebeke and Eckloff 2010).

4.4.1.2 Goffman's frame analysis

Erving Goffman's *FRAME ANALYSIS* (Goffman 1974) offers a systematic and socio-cognitively oriented theorization option for interpersonal and social interaction, especially from the perspective of sociology: instead of observing the interactions themselves, the frame (see also «Situational Context») that structures the interaction and in which the interaction is taking place is observed. In interactions the participants can negotiate and shift the frame from within the situation itself. This may create tension in the social space when the interacting agents seek to shift the frames in their respective favour (Goffman 1974, p. XV). In frame analysis, it is less a question of negotiating a social structure (in the sense of 'institution') than of experiencing interaction. *Frames* are interpretation schemas that enable the interacting agents to organise their everyday experiences and determine the definition of the situation for the interacting agents (Goffman 1974, p. 11).

4.4.1.3 Social relations model

The SOCIAL RELATIONS MODEL (SRM, Back and Kenny 2010; Kenny 1994; Kenny and La Voie 1984) offers another method for representing and quantifying social relations. It models social phenomena such as attraction, aggression, assistance, persuasion, friendship, and cooperation as interactionally dyadic and formally triadic processes (Back and Kenny 2010, p. 855). In the SOCIAL RELATIONS MODEL, perception is divided into three components: **perceiver**, **target**, and their **relationship**. As spelled out in example 6, the *perceiver effect* describes how perceivers tend to perceive other people and objects. The *target effect* describes how the target tends to be perceived by other people, and the *relationship effect* describes how the perceiver perceives the perceived individually and independently of either average perceptual tendency (Back and Kenny 2010, p. 857). These three components (also relevant to «Individual Specifics») can be used to answer various fundamental questions about interpersonal perception. According to the model, every behaviour results from a general behaviour tendency of the actor (*actor effect*), a general tendency of the partner (*partner effect*), and a specific relational action between both that is independent of the other effects (*relationship effect*).

Example 6: The social relations model according to Back and Kenny (2010)

Two people, Alan and Debby, meet at a bar. Alan judges Debby to be very intelligent. This might be because Alan perceives most people to be intelligent (*high perceiver effect*), because Debby is perceived by most people to be intelligent (*high target effect*), and/or because Alan attributes a particular intelligence to Debby that exceeds the value that he typically attributes to others and the value typically attributed to Debby by others. Alan's perception of Debby's intelligence results from a combination of these three dimensions (ibid. p. 857f.). The fondness between the two individuals can be calculated similarly: Alan might like Debby because he generally tends to like people, because Debby is generally liked by people, and because Alan likes Debby in particular.

The mathematical model of interpersonal relationship effects derived from this model offers a bridge from a human-to-human dimension of interpersonal relationship effects towards being able to model these effects on a technical level.

Examples of computer programs that analyse social relations of this type include SOREMO (Kenny 1998) and BLOCKO (Kenny 1998).

4.4.1.4 Expectancy violations model

The relational criterion ‹familiarity and relationship aspects› covers the expectations of the interacting agents. To understand the influence of expectations, and expectations of expectations on the judgement of social appropriateness, EXPECTANCY VIOLATIONS THEORY (Burgoon et al. 1988) can be utilized. This theory describes people’s reactions to the violation of expectations (for example arising from social norms). It states that unexpected behaviour triggers a cognitive-affective evaluation that can be either positive or negative. Violation leads to *arousal* in the person whose expectations were violated, then to a series of cognitive processing steps to deal with this *arousal* and to come to terms with the violation of expectations. The evaluation of violated expectations depends on various factors, e.g., the relationship with the other person or their perceived attractiveness (cf. ‹Individual Specifics›). If the violation is evaluated negatively, this has a negative effect on the judgement of the other person (lower credibility, attraction, or persuasiveness). If it is evaluated positively, it has the opposite effect. Conformity with expectations may therefore be a key factor of the configuration of interactions and their perception and judgement.

4.4.1.5 Status and power in interactions

According to Theodore Kemper’s POWER-STATUS THEORY OF EMOTIONS, the fundamental motivation and constraint behind every human behaviour (Kemper 2016, p. 5) are two constructs that should be distinguished and which connect all interpersonal interactions relationally, but not necessarily reciprocally: status and power. The **status** attributed to an interaction partner based on differences in background (e.g., friendship, reputation, group membership, profession, conformity, family ties: see also Mascarenhas et al. 2016, p. 938) determines the extent to which their wishes, needs, and interests are taken into consideration by the attributing party. **Power**, on the other hand, represents the extent to which one party can direct the behaviour of the other as desired. The *social importance* concept as used in SOCIAL IMPORTANT DYNAMICS in which the dimension of power is omitted aims to operationalize Kemper’s concept of status (see sect. 4.2.2).

4.4.1.6 Face-Saving view

When considering relationships between interacting agents, we must always consider any indirect goals being pursued in an interaction alongside the direct

goals (e.g., an appointment, maintaining a relationship through small talk, a meeting, etc.). An approach to illuminating this is offered by Brown and Levinson's FACE-SAVING VIEW (1978, 1987), part of the first wave of **politeness research**.¹⁵ In this approach, every utterance is viewed as a potentially *face threatening act* (FTA) that endangers the public self-image of the interacting agents which they must maintain to themselves and others (*face*). Two different types of *face* can be threatened, embedded within the needs structure of the actors: The *positive face*, i.e., the desire to be valued, respected, and recognized by others, and the *negative face*, i.e., the desire for self-determination and freedom of action. The model describes different strategies for dealing with FTAs (it is also relevant to the factors «Situational Context» and «Individual Specifics»). For example, two such strategies are polite behaviour: positive politeness reduces an action's threat to the *positive face* (e.g., rather than direct, personal criticism like 'you did this wrong', instead: 'We need to work on this more'). Negative politeness reduces an action's threat to the *negative face* (e.g., rather than a direct demand or request, instead: 'could you perhaps...'). The intensity of an FTA depends on the sociological variables of **social distance** (D), **power** (P) and their culturally coded **rank of imposition** (R), which subsume all other possible factors of influence, such as **status**, **authority**, ethnic **identity**, **situational factors**, etc. (ibid. p. 80). These variables are, in turn, context-dependent, and their context dependence is represented using situational factors of the speaker (S), and hearer (H). This results in the following formalization:

$$“W_x = D(S,H) + P(H,S) + R_x”(ibid. p. 76)$$

W_x symbolizes the numerical value of the intensity of an FTA, while $D(S,H)$ denotes the social distance value, $P(S,H)$ represents the power relationship between speaker and hearer, and R_x is the value of the culturally coded rank of imposition of an FTA. The values are expressed with numbers on a scale of 1 to n . The model also assigns values to the various strategies, and it is postulated that, as W_x increases, a rational actor will choose 'higher-value' strategies to execute FTAs to minimize personal risk.

¹⁵ Politeness research is an established branch of linguistics research but has long been linked to cultural studies and the social sciences through sociopragmatic approaches. For a more detailed discussion of politeness research and the *face-saving view* in the context of the phenomenon of social appropriateness, see Bellon et al. (2020, S. 274 ff).

4.4.1.7 Etiquette engine

The ETIQUETTE ENGINE (Wu et al. 2011) gives an example of an application of the FACE-SAVING VIEW to human-computer interactions. Building upon the formalization of the intensity of an FTA, Miller and others (Miller et al. 2006; Vilhjalmsson et al. 2007) designed a **politeness algorithm** that can be used to construct the behaviour of sociosensitive or socioactive agents. A core aspect of the ETIQUETTE ENGINE is its so-called *believability metric*:

$$"B_o : I_x = B_o : V(A_x) - B_o : W_x" \text{ (Wu et al. 2011, p.83)}$$

B_o symbolizes the beliefs of an observer (e.g., the hearer) regarding a perceived imbalance in an interaction x (I_x). B_o is the difference perceived by the observer between the intensity of the strategies applied in the interaction to reduce the FTA intensity ($V(A_x)$), and the face threat intensity due to x perceived by the observer (W_x). It is assumed that an interaction x is generally perceived as socially appropriate by a viewer if the intensity of the strategies used in the interaction x to reduce the FTA intensity and the face threat intensity of the FTA—as perceived by the observer—are roughly the same. If, from the perspective of the observer, the intensity of the strategies used to reduce the FTA intensity exceeds the face threat intensity of the FTA by a certain value, the interaction x is perceived to be polite; if it falls below a certain value, the interaction x is perceived to be impolite. Thus, if a person directly communicates a desire as a demand, e.g., ‘I want the salt’, this is an FTA whose intensity may vary depending on the relationship between the two individuals. One strategy to reduce the FTA is, for example, to phrase it as ‘Could you please pass me the salt?’ rather than formulating a direct demand or request. If the intensity of the applied strategy exceeds or at least matches the intensity of the desire, the interaction is perceived to be polite. Consider now a desire with a higher face threat: for example, a person might want to borrow an item of clothing from another person. Here, a mitigation strategy such as ‘Please could you lend me your scarf?’ could still be perceived as socially inappropriate, since the mitigation strategy is not sufficient to compensate for the intensity of the FTA.

4.4.1.8 The politeness principle

Another instructive commentary of social appropriateness and politeness in the context of politeness is the discussion of the so-called ‘**politeness principle**’ by Geoffrey Leech (1983). Following on from and adding to Grice’s cooperation principle and cooperation maxims (Grice 1993 [1967], p. 26–31), Leech reformulates this principle as a communication principle whose observance (especially

together with Grice's maxims) can be expected to enable (more or less) smooth communication (Leech 1983, p. 132). The politeness principle consists of **six maxims** (cf. Fig. 4.7) whose relative weighting is **culture-dependent** (ibid. p. 150):

In this context, Leech also emphasizes the **power** or **authority relationship**, as well as the **social distance** (depending on the specific **status**, **age**, **degree of familiarity**, etc.) between the actors (ibid. p. 126 f.).



Fig. 4.7 The six maxims of the politeness principle. (Figure modified after Leech 1983, p. 150)

4.4.2 «Relations between Interacting Agents» and sociosensitive/socioactive artificial assistant

The factor «Relations between Interacting Agents» should be considered highly relevant to the design of socially appropriate behaviour in interactions between humans and artificial social assistants. First, the social relationship between humans and artificial assistants is fundamentally in question: does the assistant appear as a friend, companion, possibly even a sexual partner, or simply as intelligent service equipment? This already suggests links to the frequency of the interaction, the possibility of a 'degree of familiarity' or simulated version thereof (in the sense of a *companion robot*), the respect and benevolence relationship, and the power relationship, which together constitute the criterion «familiarity and relationship aspects». The design of the robot or artificial social assistant's behaviour should therefore largely reflect how well the interacting agents know each other, how often they have previously interacted, and how they relate to one another in terms of familiarity. In turn, conclusions derived from the criterion «familiarity and relationship aspects» provide information about the criterion «intention», which describes the interests underlying the interaction. A *companion robot* has a different spectrum of potential interaction intentions than a robot that engages in a service-related interaction. Accordingly, a robot that acts solely as a service provider could be equipped with more formalized interaction behaviour than would be appropriate for a *companion robot*. The final relational criterion «context» asks in what capacity the interaction is taking place, which is closely related to the factor «Situational Context». For this criterion, the most relevant concern for an artificial social assistant is to establish a consensus about 'as what' the interaction is taking place. For the perception and judgement of social appropriateness, it might therefore be relevant to explicate the framework conditions of an interaction between humans and robots, and to incorporate more explanatory aspects into the interaction than would typically be necessary or appropriate for an interaction between humans. This will remain especially true for as long as robots and artificial social assistants are not yet fully integrated into everyday life and their use is taken for granted. In summary, the key focus of the «Relations between Interacting Agents» factor is the «familiarity and relationship aspects» criterion, although the other criteria cannot be neglected.

4.5 The «Standards of Customary Practice» factor

The final factor of the construction, perception, and/or evaluation of social appropriateness in interpersonal interactions concerns the action and behaviour-regulating «Standards of Customary Practice» that apply to a specific action or

behaviour, understood here as an ‘**ensemble of the usual**’.¹⁶ This factor encompasses situational, group-related, and general maxims, and/or uncircumventably established concepts (e.g., ‘For our volleyball team, being up to 15 min late to training is tolerable’) and social standards, including moral norms (moral standards) that apply to every party¹⁷ (e.g., protection of dignity), which therefore also claim to be **ethically** justifiable. Different ensembles of the usual can be understood on a spectrum ranging from ‘**specific**’ (and possibly divergent) **group morals** (e.g., differences in family morals, sports team morals, or company morals) on the one hand, which in extreme cases are constantly situationally renegotiated, to ethically justifiable, and hence ‘**general**’, **regulative norms** on the other.

Possible factor criteria:

- **«Habitus»:** What are the usual and ‘ingrained’ types of behaviour and judgement of an individual?
- **«Values/social norms»:** According to which values/social norms—from group-related norms to societal institutions—is social appropriateness judged? Including institutionalized values and institutionalized social norms.
- **«Regulative norms»:** What social norms can be ethically justified?

Examples of questions and aspects that can be assigned to each factor criterion to support an analysis of the potential social (in)appropriateness of interaction situations and guide these situations are given in Table 4.5.

Possible observables:

- Voting behaviour; party donations; eye contact; loudness; professional attire; accessories; clothing style; nutrition, etc. as general observables of the habitus relating to ingrained typical types of behaviour in the context of social appropriateness, such as: what physical distance is perceived to be appropriate in which types of interaction situation? How do you greet someone, say goodbye, etc.?
- Legal texts, ordinances, protocols, etc.; judging behaviour or actions, including group sanctions where necessary (sometimes explicitly stated: ‘Anyone who gets a yellow card for arguing with the referee pays €5 into the team jar’);

¹⁶This usage takes up that of Odo Marquard, cf. e.g., Marquard (1979).

¹⁷Which does not necessarily imply that the norm is unconditional. Even the protection of dignity should be understood as a conditional norm that does not need to be followed when acting in defence in a life-or-death situation. Thus, it applies to everyone, but not always.

Table 4.5 Standards of Customary Practice. (Source: own representation)

«Standards of Customary Practice»	
«Habitus» What are the usual and ‘ingrained’ types of behaviour and judgement of an individual?	What conduct, demeanour, etc. does a person display in the context of social appropriateness? For example: how do they behave at the table, how do they greet someone, and what do they talk about?
	How do they appear, how do they present themselves, if applicable? For example: what style of clothing do they wear, how do they talk about themselves?
	What job and what activities do they pursue?
«Values/social norms» According to which values/social norms—from group-related norms to societal institutions—is social appropriateness judged? Including institutionalized values and institutionalized social norms	What expectations and rule systems are there in a group (culture, subculture, association, circle of friends, but possibly also random groups like ‘people currently shopping at IKEA’) in the context of social appropriateness? What typical practices have developed within this group and other similar group settings?
	What institutionalized social norms are there? For example: what legal framework is generally applicable? Are there any special or area-specific rights?
	What counts as violating a framework setting (negative determination of appropriateness)?
	How are violations of boundaries institutionalized? For example, what is legally defined as violence, and what is ‘merely’ socially defined as emotional overstep?
	How are the possibly institutionalized values and social norms indicated? (statutes, etiquette handbook, rules of conversation, company compliance, constitution, laws, police legislation, social sanctions such as ‘stern looks’, ‘raising the eyebrows’, praise, rewards, etc.)
«Regulative norms» What social norms can be ethically justified?	Is there an ethics council whose opinion can be sought for certain topics? Have certain ethical positions already been established? (e.g., by ethics councils in businesses, or at state level)
	What is the state of ethics research on the ethical problem being debated?
	Is a specific ethics case study necessary?
	Regarding research: what so-called <i>ELSI</i> or <i>ELSA</i> (<i>Ethical, Legal, Social Implications/Aspects</i>) are involved, if any? How should they be handled? Should they be decoupled or addressed by <i>integrated research</i> (Gransche and Manzeschke 2020)?

concrete reactions to a behaviour or action (praise, reprimands, facial expressions, gestures, etc.), and so on.

- State of the art of research in ethics; results of ethical case studies; opinions from various ethics councils.

4.5.1 Theoretical perspectives on «Standards of Customary Practice»

Besides **regulative norms**, the concept of **social norms** plays a central role in the connection between the phenomenon of social appropriateness and «Standards of Customary Practice». This distinction between social norms and regulative norms also reflects the distinction between the meaning structures that organise people's shared lives (norms from a descriptive perspective) and general action orientations (norms from a prescriptive perspective). Regulative norms become ethical norms precisely when their claim of justification can be redeemed.¹⁸ Social norms have been examined from very different perspectives and play a central role in various disciplines. For a practicable and interdisciplinary understanding of the concept of social norm within the scope of this book, we could for example follow the position by Krettenauer and Montada (2005) that norms fundamentally regulate coexistence within social communities, and can essentially be understood as "**sanction-enforced behavioural regularities**" (Diekmann 2022).¹⁹ This can take the form of imperatives and prohibitions, duties and responsibilities

¹⁸Regarding this basic distinction in the concept of norms, see also the glossary entry and Ganslandt and Wimmer (2005); Kambartel (2005). Since ethical considerations regarding social appropriateness and artificial assistants are discussed in Chapter 7 (among other things), they are not examined in any further detail here. Regarding the connection between social appropriateness and regulative norms of ethics, it should however be noted that, from a purely descriptive perspective of norms (in connection with «Standards of Customary Practice»), violations of regulative norms of ethics can be considered socially appropriate (consider discriminatory practices in various historical contexts). But including an ethical perspective into the phenomenon of social appropriateness explicitly prevents this; in other words, in the overarching understanding of social appropriateness adopted here, it is assumed that any actions that violate regulative norms of ethics are socially inappropriate. (Although there is some overlap, the fact that socially inappropriate behaviour may be ethically necessary is a separate question.).

¹⁹For the concept of social norm see also Hechter and Opp (2001); Opp (1983), and Bicchieri (2006, 2017).

towards other people, groups, and institutions, or specific rights. Social norms are sometimes expressed in laws, customs, or conventions and are primarily characterized by the fact that views about their legitimacy and functionality differ. For example, they might be rejected in one community or culture but considered proper in another. The phenomenon of social appropriateness is closely linked to social norms: what is considered socially appropriate in a specific interpersonal interaction situation also depends on the **prevailing socially appropriate ‘standards of customary practice’**. These standards act as rules of action for the interacting agents, applicable not only to a particular interaction situation, but as **social norms** for any sufficiently similar situations.

4.5.1.1 Approaches to «social norms» based on theories of game rules

In the context of social appropriateness, game rule theories state that the (in)appropriateness of social behaviour is (at least partly) determined by constitutive rules, similar to the rulebook of a game (e.g., a board game, the classical philosophical example being chess).

A prominent variant of these approaches is represented by John Rogers Searle. For Searle, social issues consist of constitutive rules of the form **X is considered Y in context K** (with K also referring to the factor «Situational Context») and are conveyed through the processes of socialization and habitualization (Searle 2002). For Searle, constitutive rules also establish normative standards of appropriateness for the realization of social practices:

[...] there is a socially created normative component in the institutional structure, and this is accounted for only by the fact that the institutional structure is a structure of rules, and the actual rules that we specify in describing the institution will determine those aspects under which the system is normative. (Searle 1995, pp. 146–147)

According to Searle, the constitutive game rules of the community therefore have the function of determining which behaviour is considered inappropriate/appropriate—in which contexts. Regarding the inappropriateness/appropriateness of interpersonal interactions, the corresponding constitutive (overarching) rule thus therefore the following form:

Behaviour X is considered inappropriate/appropriate in context K.

A second group of approaches based on theories of game rules are called ‘practice theories’ and postulate that generally binding rules are constitutive for social

practices—for example, Rahel Jaeggi writes: “Practices are *rule-based*. [...] They always concern courses of action that are determined by rules and regulations, from a classification of action alternatives into what is appropriate to do and what is not.” (Jaeggi 2014, p. 97) The rules immanent in a group’s social practices therefore ultimately determine whether interpersonal behaviour is considered appropriate or inappropriate.

This approach, and similar approaches based on theories of game rules are the subject of debate within philosophy. As an example of a philosophical approach that marks a critical departure from game rule theories, consider the observation that there is a lack of objective standards of appropriateness in cultural practices; generally binding rules for properly implementing appropriateness standards can be understood as negotiable in any given situation (Poljanšek 2019). Standards of social appropriateness therefore are constantly being reinvented by social practice. The intelligibility and rough stability of standards of social appropriateness appear to ultimately reflect similar rule expectations based on shared actualities.

4.5.1.2 Rational choice theory and game theory

Another theoretical approach to **explaining social norms** is provided by so-called RATIONAL CHOICE THEORY (Coleman and Fararo 1992), which exerts a powerful influence in multiple disciplines (and is also relevant to the factor «Situational Context»). In essence, *rational choice* approaches argue that social norms arise and exist if and because following them is rationally necessary, in the sense that they contribute sufficiently to maximizing individual utility. This theory is based on the idea that all human actors are fundamentally rational, and that both the costs and utilities are ‘calculated’ in advance during decision-making. According to this model, people follow utility-maximizing or cost-minimizing principles and have different interindividual behavioural preferences (cf. «Individual Specifics») motivated by individual goals and wishes. Different **framing** in the communication of information leads to different decision-making behaviour. The theory proposes that people always choose the alternative that offers the greatest satisfaction (in the sense of maximum utility and minimum cost). This framework also offers a plausible explanation of violations of the social norms of appropriateness: for example, behaviour considered socially inappropriate (e.g., throwing chewing gum on the floor) could represent the cost-minimizing alternative in the evaluation of the offender despite requiring a norm to be violated, since the effort of looking for a bin to dispose of the chewing gum is perceived to be associated with a higher cost than the violation of the norm (possible disapproval from peers, etc.).

Rational Choice Theory plays a role in game theoretical modelling of social interactions (Diekmann 2016; Voss 2001). GAME THEORY assumes that ‘players’ in model games act goal-oriented to receive the highest possible pay-out. Game theoretical models can be used to illustrate various types of social norms such as coordination norms, cooperation norms, and distribution norms (Diekmann 2019, 2022) based on a differentiation of social norms in terms of their contribution to conflict resolution in interactions (“coordination norms”, “PD norms”, and “norms of partiality”) established by Edna Ullmann-Margalit (Ullmann-Margalit 1977):

- **Coordination norms** select one of multiple so-called ‘Nash equilibria’. In a game, a Nash equilibrium is a particular configuration of strategies in which each player chooses a strategy, and it is not rational for either player to be the only one to deviate from their choice. In a Nash equilibrium, each player is therefore content with their choice of strategy, even retrospectively—the strategies of each player are accordingly the best possible response in each case. A frequently cited example of a coordination norm is driving on the right (countries in continental Europe) and driving on the left (Great Britain). Two drivers approaching one another can both choose either ‘left’ or ‘right’, resulting in four combinations, two of which produce Nash equilibria and the other two of which lead to collisions that damage both drivers. The requirement to drive on either the left or the right then corresponds to selecting one of these two Nash equilibria. A similar principle can be applied to handshakes as a greeting ritual: social norms define the ‘greeting hand’ (left or right). Of course, in this case a ‘collision’ would typically not prove life-threatening, but nor would it be useful (i.e., suitable for the purpose of greeting). Extending one’s ‘right hand’ is—under ‘normal conditions’ unaffected by the coronavirus—a coordination norm. Thus, this norm regulates what behaviour is **appropriate** in certain cultural settings during greetings, and some **socially appropriate standards of customary practice can be understood more precisely as coordination norms**.
- **Cooperation norms**, on the other hand, fall under the category of collective goods, i.e., goods from which everyone benefits, even without personally contributing as a so-called ‘free rider’. Examples of collective goods for example include a well-preserved environment or adequately funded public transport. Some typical cooperation norms relating to these examples are ‘don’t throw rubbish onto the street’ or ‘don’t be a fare-dodger’. In a fictitious state with no applicable norms, the use of collective goods could initially be formulated as a so-called ‘prisoner’s dilemma’, essentially: rather than mutual cooperation,

mutual non-cooperation is a Nash equilibrium; in the prisoner's dilemma, each prisoner fares best when they are the only one to choose not to cooperate individually—'fare-dodgers' don't need to pay for transport because the other passengers pay for them.²⁰ However, if everyone follows the same logic, the overall result is 'inefficient' (or 'not Pareto-optimal')—if nobody pays the fare, public transport cannot be financed in the first place. The solution to this situation is a 'contract' stabilized by sanctions: everyone promises to cooperate—e.g., by paying the fares—and are sanctioned if they deviate from this promise. The corresponding social norms are the so-called 'cooperation norms'.²¹ Some **socially appropriate standards of customary practice**, such as not throwing rubbish onto the street in the above example, **can therefore be more precisely understood as cooperation norms.**

- Finally, **distribution norms** regulate the distribution of goods of all kinds, not just money, but also services or waiting times in queues. In most cultures, distribution norms fall under the regulative norm of egalitarian treatment, i.e., sufficiently justified distribution inequality. A distribution norm that is straightforward but functional even for rational egoists in this sense is the principle of 'one person divides up and the other person chooses'. Here, a Nash equilibrium strategy is to always give up exactly half because any deviation from equal sharing would be punished by the other player. Among other things, this social norm has been used to establish appropriate settlements of divorce and inheritance disputes (Brams and Taylor 1996), e.g., one of the parties places red or blue stickers on all items in the previously shared household. After inspecting the stickers, the other party in the dispute then chooses a colour and received all objects marked with the corresponding stickers. As this special case of 'appropriate divorce behaviour' demonstrates, **some socially appropriate standards of customary practice can also be reconstructed as distribution norms.** The same applies to waiting times for queues: the fact that it is socially appropriate not to push ahead in a queue can be understood as a distribution norm that governs the waiting times of the people in the queue.

²⁰This is of course a simplified example that disregards the fact that local public transport may be financed by taxes to a certain extent, as is the case in Germany.

²¹Any 'second-order free rider problems' (Coleman 1990; Heckathorn 1989), which arise when the sanctions are themselves associated with costs and consequently neglected, leading to the overall collapse of cooperation as a result, are not considered here.

Coordination, cooperation, and distribution norms typically do not manifest in a ‘pure’ and ‘idealized’ form, but present in various mixed forms that often involve coordination problems. This is also the case for many socially appropriate standards of customary practice, where resolving distribution conflicts plays a role alongside questions of coordination. This is illustrated by the following example 7, based on the so-called ‘*chicken game*’ from game theory²²:

EXAMPLE 7 Chicken Game

A couple wants to enter a hotel through a revolving door; one person must go first. If both go first (...) there is a collision. If both wait (...) time is lost until someone ultimately decides to go first. There are two ways to pass through the door without wasting time: the wife goes first (...) or the husband goes first. Both “strategy profiles” are equilibria (...). The social norm of “the man allows the woman to go first” ultimately amounts to selecting an equilibrium to ensure an “efficient” interaction. (cf. Diekmann 2022, own translation)

If both agents have the same gender, the corresponding customary standard of social appropriateness can instead be based on the criterion ‘age (of service)’ (principle of seniority).

Finally, another category of social norms especially relevant to the socially appropriate standards of customary practice that captures the meaning structures that organise people’s shared lives, which cannot always be understood from the perspective of a rapid cost-benefit calculation, is that of so-called ‘**signal norms**’ (Diekmann 2019, 2022). The associated fundamental *SIGNALING THEORY* can be traced back to Spence (Spence 1973), who used it in an attempt to explain investment in education (as a signal to an employer), and Zahavi (Zahavi 1975), who used it to explain biological phenomena (peacock feathers as a signal of health). Posner (Posner 2000) proposed the first systematic connection to social norms. Further discussions of signal norms focus in particular on cost-intensive

²²In the original scenario of the chicken game, two cars are driving towards one another. The first car to swerve to one side is the ‘*chicken*’, a coward who loses the game. If neither car swerves, the cars will collide. Mutually deciding not to swerve is not a Nash equilibrium—unlike in the prisoner dilemma—because there is always an incentive to do so while the other player remains on a collision course.

signal norms: whoever follows a signal norm signals a willingness to cooperate; whoever violates it is presumably uninterested in repeated interactions and can be expected not to cooperate (see also Diekmann and Przepiorka 2010 and the experiment by Przepiorka and Diekmann 2013). The typical example of a signal norm is an engagement gift (see also Camerer 1988)—although this example will not seem equally compelling to everyone; the context of this example is American wedding culture: inexpensive engagement gifts expose potentially unreliable husbands, who in extreme cases may not even be seriously interested at all. Signal norms stabilize human interactions by signalling a willingness to cooperate—in different senses and at different levels: anyone who buys *stone-washed* jeans that are more expensive than jeans designed without a *used look* is not acting irrationally but displaying a (usually subconscious) willingness to follow certain fashionable trends that act as a stabilizing factor in interactions with similarly minded people and possibly a way to open interactions. **Many of the socially appropriate standards of customary practice, in court ceremonies, dining and table manners, fashion standards, or physical forms of expression can be (alternatively) reconstructed in this sense as signal norms** (Diekmann 2022).

4.5.1.3 Norm activation model

To explain sets of social appropriateness customs, we must also explain how social norms could possibly **influence behaviour** («Type of Action, Conduct, Behaviour, or Task») in the first place. For a norm to be considered to guide behaviour, it must be visible/recognizable or internally activated for a given individual («Individual Specifics»). Alongside other theories, the so-called **NORM ACTIVATION MODEL** (see Fig. 4.8) describes under which circumstances norms become salient, or in other words are activated in individuals, and under which circumstances they influence behaviour (Schwartz 1977; Schwartz and Howard 1981). According to this model, personal norms are actively experienced as a feeling of moral obligation rather than any purely behavioural intention. They are influenced by the sense that performing (or failing to perform) a certain action may have consequences, as well as a sense of responsibility to display this action to others. By internalizing norms, people perceive the norm as belonging to them, which causes deviant behaviour to generate feelings of guilt.

To activate a norm, various necessary steps must be performed. (Schwartz (1977) and Schwartz and Howard (1981).)

The steps necessary to initiate an action are summarized as follows: a person becomes aware of a problematic circumstance that requires rectification. Responsibility is attributed for this negative circumstance (to institutions/people/oneself).

I. Activation steps: perception of need and responsibility

- 1. Awareness of a person in a state of need
- 2. Perception that there are actions which could relieve the need
- 3. Recognition of own ability to provide relief
- 4. Apprehension of some responsibility to become involved

II. Obligation step: norm construction and generation of feelings or moral obligation

- 5. Activation of preexisting or situationally constructed personal norms

III. Defense steps: assessment, evaluation, and reassessment of potential responses

- 6. Assessment of costs and evaluation of probable outcomes (The next steps may be skipped if a particular response clearly optimizes the balance of costs evaluated in step 6. If not, there will be one or more iterations through steps 7 and 8.)
- 7. Reassessment and redefinition of the situation by denial of:
 - a) state of need (its reality, seriousness)
 - b) responsibility to respond
 - c) suitability of norms activated thus far and/or others
- 8. Iterations of earlier steps in light of reassessments

IV. Response step

- 9. Action or inaction response

Fig. 4.8 The norm activation model. (Modified after Schwartz 1977 and Howard 1981)

There are perceived ways to alleviate or resolve the relevant circumstance, and the ability to act accordingly is also perceived to exist.

4.5.1.4 Manners

To illustrate different theoretical approaches in the context of «Standards of Customary Practice» and «social norms», it is also enlightening to look at a sub-form of social norms that has been studied specifically in philosophy and sociology: **manners** (that depend on the factors «Relations between Interacting Agents», «Situational Context» and «Individual Specifics» and influence the «Type of Action, Conduct, Behaviour, or Task»). The fact that interpersonal interactions are shaped by certain manners is a well-known phenomenon in everyday life: you wouldn't spit at anyone you weren't trying to insult. Similarly, you would not expect to be asked detailed questions about your personal sex life in public, and you would not attempt to get laughs out of anybody after news of a bereavement in the family. Especially when considering the possibility of transferring standards of socially appropriate behaviour between people to interactions between people and artificial assistants, we must ask the fundamental question of what functions interpersonal manners ultimately fulfil in people. Commentaries by

Immanuel Kant, Norbert Elias, Erving Goffman, and Pierre Bourdieu are outlined below as an example.

For Immanuel Kant, the function of certain interpersonal manners - or **forms of propriety** (*Anstandsformen*) - is the **moral education of man** (Gransche 2019; Paret 2019): forms of propriety indicate a moral disposition that can eventually lead people who repeatedly behave **with propriety** to **moral decency and morality**. The title ‘On permissible moral illusion’ of a chapter of Kant’s *Anthropology from a pragmatic point of view* (Kant 1983 Erstveröffentlichung: 1798), which presents the Kantian theory of propriety, already indicates that Kant views forms of propriety as essential forms of human practice. For Kant, these forms of propriety do not necessarily require an authentic moral disposition or genuine moral decency, that is: not always and not to begin with, but the mere forms of propriety can lead to moral decency. Indeed, **propriety in general merely represents the illusion of a respective moral disposition, but this is not morally problematic**. He expresses this especially clearly in the following passage:

In order to save virtue, or at least lead the human being to it, nature has wisely implanted in him the tendency to allow himself willingly to be deceived. Good, honorable *decorum* is an external illusion that instils *respect* in others [...] In general, everything that is called *propriety* (*decorum*) is of this same sort – namely nothing but *beautiful illusion*.

Politeness (*politesse*) is an illusion of affability that inspires love. *Bowing* (compliments) and all *courtly* gallantry together with the warmest verbal assurances of friendship are to be sure not exactly always truthful (“My dear friends: there is no such thing as a friend.” *Aristotle*); but this is precisely why they do not *deceive*, because everyone knows how they should be taken, and especially because these signs of benevolence and respect, though empty at first, gradually lead to real dispositions of this sort. (Kant 2006, p. 43-44)

In his two-volume work *The Civilizing Process: Sociogenetic and Psychogenetic Investigations*, German sociologist Norbert Elias addresses the **origins of types of behaviour** deemed to be **civilized** and therefore ultimately the origins of ‘civilized manners’. Elias’ investigation specifically considers the traditional standards of behaviour passed down from the ninth to eighteenth century in the ‘Occident’, explicitly reporting developments in Germany, France, Italy, and England, in particular **table customs**, such as the following rules taken from Tannhäuser (thirteenth century):

A man of refinement should not slurp from the same spoon with someone else; that is the way to behave for people at court who are often confronted with unrefined conduct.

It is not polite to drink from the dish, although some who approve of this rude habit insolently pick up the dish and pour it down as if they were mad.

Those who fall upon the dishes like swine while eating, snorting disgustingly and smacking their lips...

Some people bite a slice and then dunk it in the dish in a coarse way, refined people reject such bad manners.

A number of people gnaw a bone and then put it back in the dish – this is a serious offence. Elias 2000, p. 73

In essence, for Elias, the intertwined origins of civilized and implicitly or explicitly codified rules of conduct and interpersonal manners are ultimately the functional differentiation of society based on the division of labour and the growing individual affect control of humans. With this analysis, Elias steps away from a strongly individual rationalization of the development of ‘civilized’ interpersonal types of behaviour and interaction: the civilizing process is driven not by insight, for example that certain behaviours like washing one’s hands after relieving oneself minimizes the transmission of disease, but instead by interaction with individual affective states and the social structuring of society. For Elias, the civilizing process unfolds from top to bottom, that is, starting from the nobility and slowly spreading throughout the population. For an understanding of the phenomenon of social appropriateness in interpersonal interactions, the reflections of Norbert Elias illustrate the **central role of social appropriateness within human development** and the central role of **societal and psychological structures** in the context of the development of socially appropriate standards of customary practice.²³

For Erving Goffman (Goffman 1982, 2009), on the other hand, manners, which he describes as being codified into so-called ‘**ceremonial rules**’, serve to **socialize** individuals without excessive imposition on their personal convictions.²⁴ They have a purely performative character, can be fulfilled by as many people as possible with ease, thereby enabling **social integration**, and offer the possibility of individual **sovereignty** in a strong sense: here, sovereignty does not just mean that the rules associated with manners are easy to follow. Manners consist of a

²³ For reflections on the capability of acting socially appropriate as a cultural technique referring also to this central role of social appropriateness for human development, cf. Nähr-Wagener (2022).

²⁴ The connection between social appropriateness and Goffman’s considerations is discussed in depth in Christoph Paret’s research report for poliTE, cf. Paret (2019).

myriad of seeming trivialities (How tightly should you squeeze a person's hand when greeting them? How much should you smile at a compliment? What facial expressions should you have when praising someone?) that are at the same time more important than anything else (Goffman 1982c, p. 503)—they demonstrate an effortless mastery of the situation. By “returning a greeting when greeted”, a person shows that “they are sufficiently in control of their current situation to be able to spare a moment for greetings [...] friends who notice that a greeting was not returned will have the impression that something is wrong” (Goffman 1982b, p. 113). According to Goffman, through effortless and smooth mastery of manners, people demonstrate that they are in a mode of **sovereign everyday mastery**, so to speak. For Goffman, human manners thus have three crucial functions:

1. Due to their purely performative character, they enable **socialization** without the risk of the personal convictions of an individual being too strongly affected by excessively strict rules of social behaviour.
2. Due to their low threshold, they enable the **social integration** of as many individuals as possible.
3. They facilitate the impression and expression of **sovereignty**.

Finally, according to Pierre Bourdieu (cf. Bourdieu 1982), manners have (at least) one other function: they serve as means in a classification struggle involving the definition, marking, and reproduction of specific **class differences**. Manners (among other factors) constitute, signal, and reinforce **class affiliation**, they are a “symbolic expression of class position” (cf. *ibid.*, p. 284, own translation) and part of the **habitus**. For Bourdieu, there are no **neutral** manners. Manners are always a normative category imposed by the so-called ‘bourgeoisie’—in Bourdieu’s work, this concept is somewhat detached from Marx’s criteria, i.e., it is not restricted to the narrow sense of the owners of the means of production. It is however not detached from the economic **basis** of its definition after Marx; in essence, the concept of bourgeoisie can be understood here in the broad sense as the social upper class. By contrast, according to Bourdieu, the lower classes do not have any other manners in a strict sense, in fact generally they have none; they typically just represent the **negative film** of the manners of the bourgeoisie. For Bourdieu, besides their status as a means of distinction, the manners of the bourgeoisie are characterized by a rejection of substantiality and a leaning towards *empty forms* such as asceticism, self-restraint, and even self-denial.

4.5.2 «Standards of Customary Practice» and sociosensitive/socioactive artificial assistants

Since the factor «Standards of Customary Practice» addresses areas of social appropriateness that have a high and possibly sensitive impact on human living, it must be given careful attention when designing socially appropriate behaviour for artificial social assistants. Special attention should be paid to the distinction between ‘concrete’ **group morals** (encompassed by the ‘habitus’ criterion, which covers ‘ingrained’ types of behaviour and judgement, and the ‘social norms and values’ criterion, which ranges up to societally institutionalized norms) that can be negotiated situationally on the one hand, and ethically justifiable—‘general’—**regulative norms** (formulated as a separate factor criterion) on the other. The severity of compliance or non-compliance with these norms varies accordingly, from indifference to different degrees of disapproval, including ‘self-reproaches’ if necessary, or even institutional sanctions.

Therefore, addressing different aspects for example in a robot’s behavioural repertoire can be accorded different levels of significance. **Institutionalized standards defined within legislative frameworks must always be considered and observed.** In worst-case scenarios, non-compliance may lead to the robot being shut down or even serious injury to human interaction partners, and negative legal consequences for the developers and organisation responsible for development. **The same applies to other ethical norms** that have not been institutionally enshrined in law (cf. also Chapter 7). Non-compliance with non-institutionalized but applicable social norms, for example fairness, may lead to disapproval or interaction termination. However, some social norms that operate within the habitus, such as differences in table manners in different social milieus for example, will only represent a necessary criterion for successful human-machine interaction in highly specialized and possibly highly ‘personalized’ contexts (imagine for example a fictitious ‘manners robot’ used for diplomatic training), but will not have to be taken into consideration in other contexts. Still, considering the ingrained customs within a group helps robot or other technical systems behaviour to be perceived as socially appropriate if they are intended to be viewed as a member of this group—what will be perceived as socially appropriate may also depend on the platform the system uses: humans may have varying expectations according to a technical system’s form of embodiment. Failure to observe customs may lead to a spectrum of reactions, ranging from complete indifference to moderate disapproval (within the relevant group). In the case of practices regarded by a group as particularly important, however, the consequences could extend to termination of an interaction or stronger disapproval,

possibly even expressed as violence against the technical system. A possible **rule of thumb** could be that ethical norms, institutionalized norms, and non-institutionalized norms that fundamentally play a central role in maintaining the interaction should always be taken into consideration, while finer-grained group customs are ‘only’ advisable to consider in specific application contexts of socially appropriate human-machine interactions.

Another important question is the **fundamental transferability of «Standards of Customary Practice»** in interpersonal interactions to human-machine interactions. It is probable that humans adapt to technical systems logic as well as follow their own logic and that machines are designed to adapt to human logic as well as perform their own. In this interplay, perhaps new rules of conduct that have not been there before can emerge for both, humans and machines. Useem and Useem (1967) proposed the notion of a ‘third culture’ that is created where, for example, a child grows up in a family that has moved to another cultural region. The child may have the experience of being offered different sets of cultural rules and «Standards of Customary Practice» in their family and in joining activities outside of their home. The child may therefore end up building and living in a so-called ‘third culture’, incorporating and acting according to potentially conflicting and/or enriching elements of both sets of standards. Regarding human-machine interaction the concept of ‘third cultures’ could be interesting: maybe new standards of social appropriateness will emerge due to the differences in human and machine logic? How do human standards change due to machine logic? How is machine ‘behaviour’ designed to adapt to human logic? Are there misunderstandings? How does culture itself change due to new technical systems being invented? Where do elements of all these areas come together to build something new? Answering these questions is subject of current and may be subject of future research.

However, based on the above, we can draw one boundary for the discussion of socially appropriate technical systems, even with appropriateness standards that were not specifically developed for human-machine interactions. In his reflections on interpersonal manners, Kant remarked that humans often **only give the illusion** of following moral attitudes by acting in certain ways, but do not necessarily hold the corresponding moral beliefs. If the actual moral attitudes of the interacting agents are irrelevant or secondary in this sense even in interpersonal interactions, then it would also be unproblematic to implement standards of social appropriateness in artificial assistants that cannot adopt their own moral positions. From this point of view, debating ontological differences and the possibility or impossibility of overcoming them between humans and technical systems is of secondary importance for the question of sociosensitive and socioactive assistance technology.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Social Appropriateness Tree Structures

5

As another way of approaching the phenomenon of social appropriateness, the factors discussed in the explanations of the *FASA MODEL*, their criteria, and the observables/indicators of social appropriateness prominently associated with them can be graphically arranged into tree graphs to illustrate their relations among themselves.

This allows rapid access to a preliminary understanding of which factors and factor criteria of social appropriateness should be considered when building systems whose design is technically limited by only being able to detect certain observables with sensors.

For example: if you want to design a technical system that ‘only’ has the ability to recognize movement data, you can check for any relevant features in the indicator section of the tree structures below and decide which factors and factor criteria of social appropriateness might be relevant. This *bottom-up approach* can also be reversed to obtain a *top-down approach*:

If it is unclear which observables need to be technically detected in the design process of a sociosensitive and/or socioactive system, the factors and factor criteria of social appropriateness can be used to assess which observables might represent potential indicators of the social appropriateness relationships involved in the application.

The tree structures also graphically show how the factor criteria <time> and <intention> are associated with two different factors in each case – for <time>, the two factors <Situational Context> and <Type of Action, Conduct, Behaviour, or Task>; for <intention>, the two factors <Type of Action, Conduct, Behaviour, or Task> and <Relations between Interacting Agents>. This illustrates the **interconnections**

between the factor criteria and the factors and therefore between the factors themselves. The factors, factor criteria, and observables/indicators of social appropriateness are represented in Fig. 5.1, 5.2, 5.3, 5.4 and 5.5 in tree structures, starting from a factor in each case.

a) «Type of Action, Conduct, Behaviour, or Task»

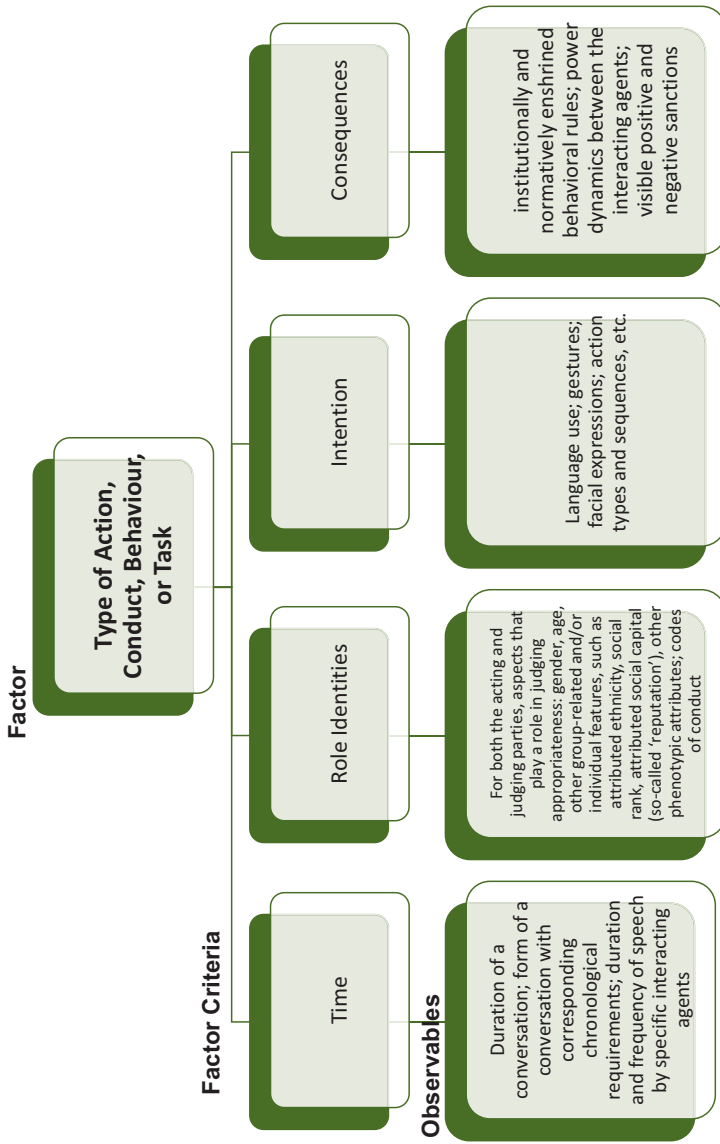


Fig. 5.1 «Type of Action, Conduct, Behaviour, or Task». Source: own representation

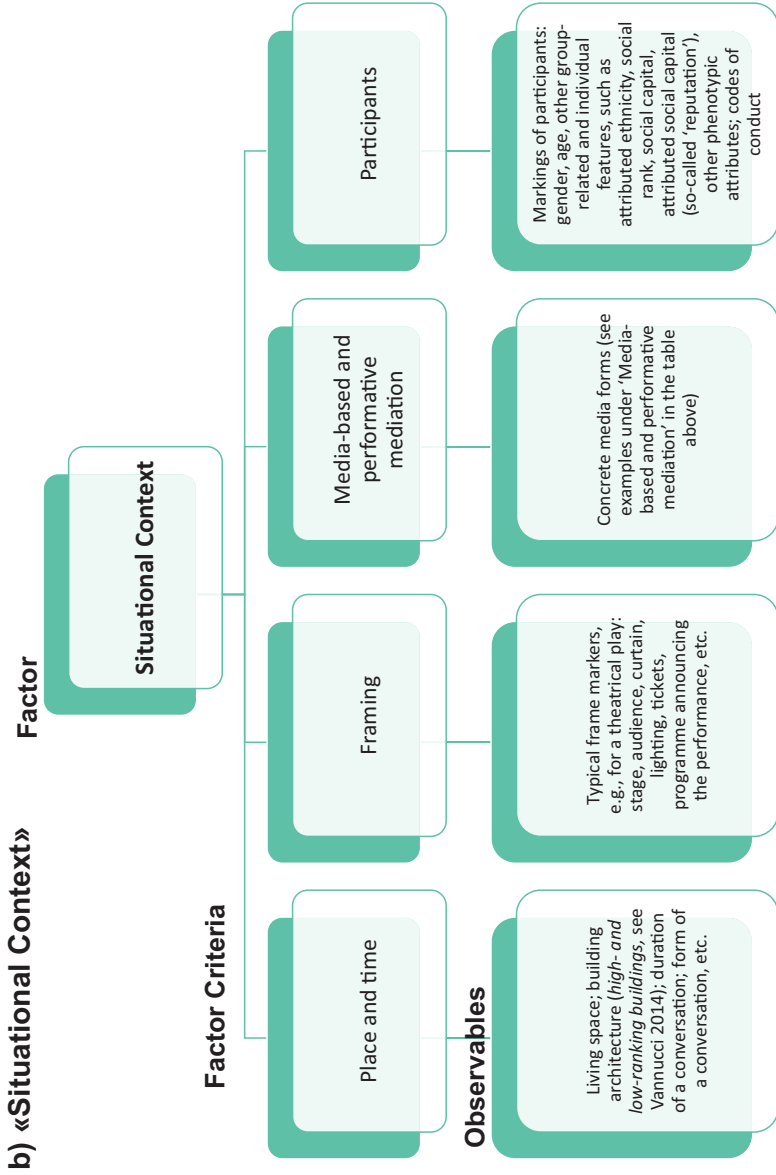


Fig. 5.2 «Situational Context». *Source* own representation

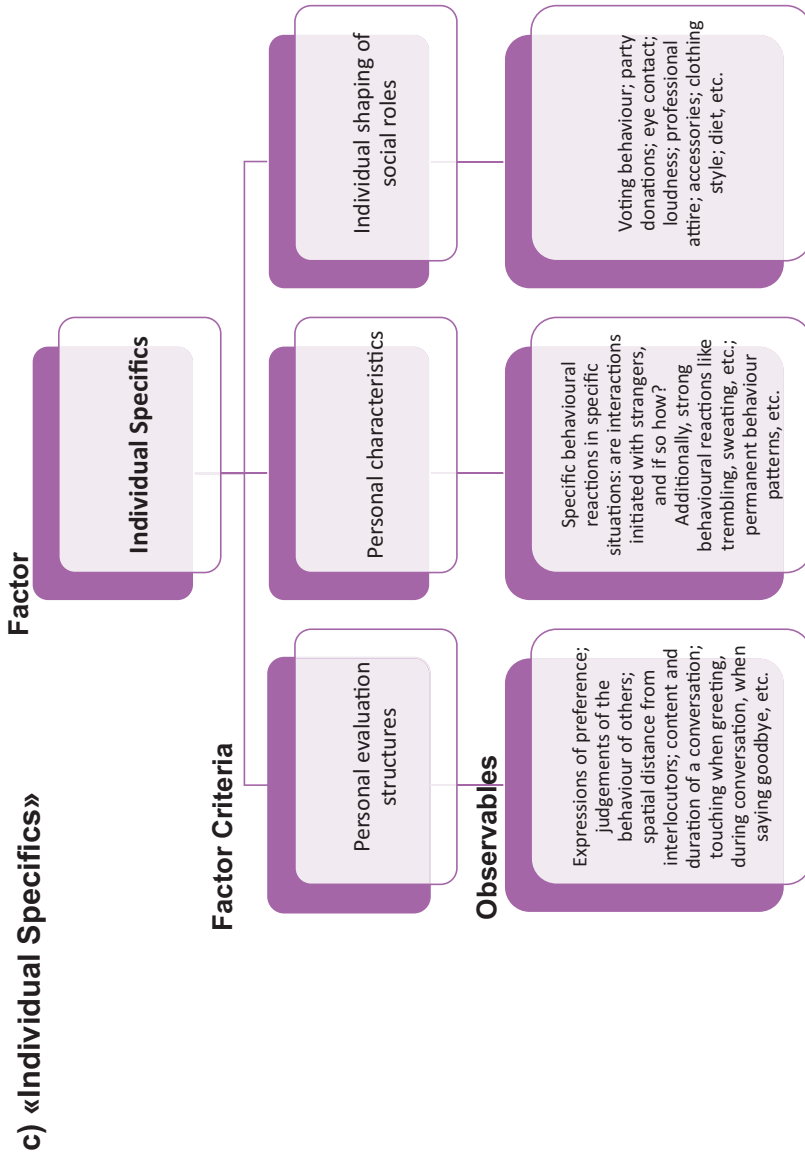


Fig. 5.3 «Individual Specifics». *Source* own representation

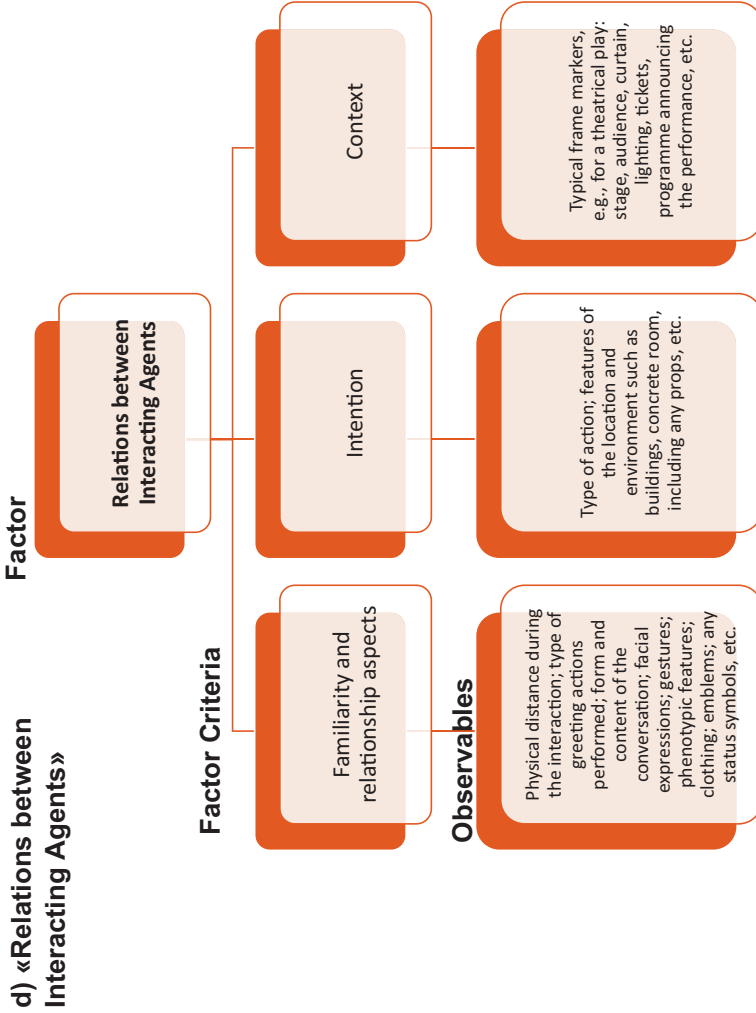


Fig. 5.4 «Relations between Interacting Agents». *Source* own representation

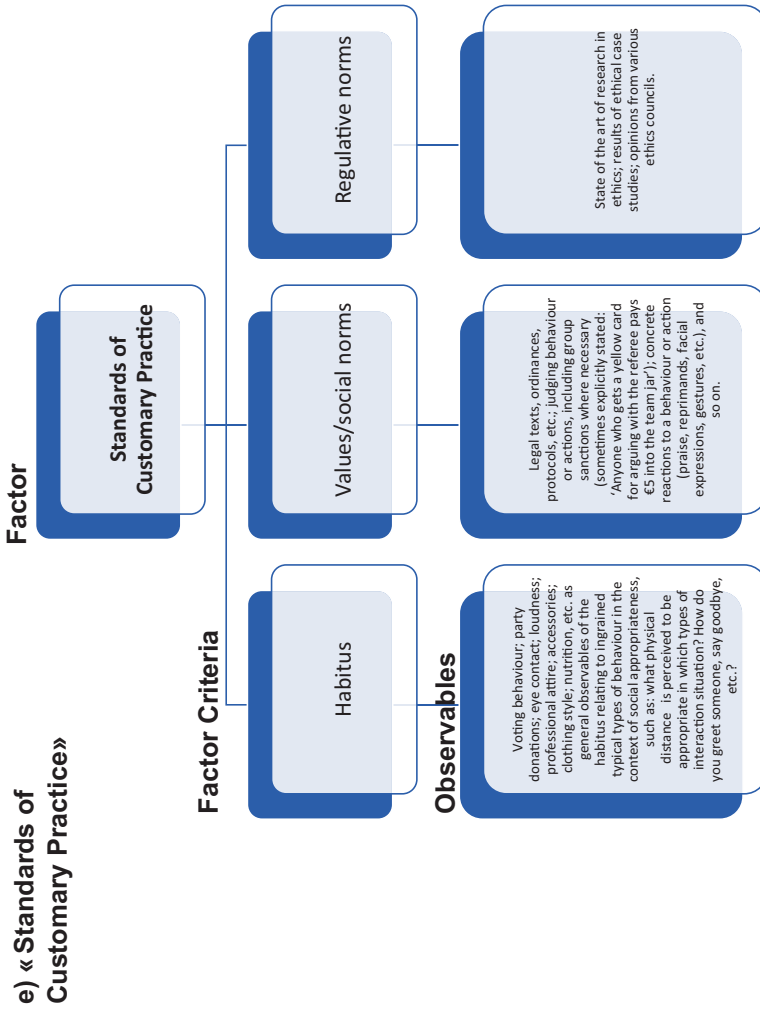


Fig. 5.5 «Standards of Customary Practice». *Source* own representation

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Application of the FASA Model to Technical Systems

6

To decide which social abilities are required for a robot and which aspects do not require so much weight, the fields of application and the type and frequency of human contact should be analysed in depth (see also the explanations for the «Relations between Interacting Agents» factor). The evaluation criteria proposed by Dautenhahn (2007), each defined on a spectrum, can be used as a basis to analyse contact within a specific field of application (Dautenhahn 2007, p. 683):

1. **Contact with humans** (none; distant to repeated; long-term; physical),
2. **A robot's functionality** (limited and clearly defined to openly adaptable; shaped by learning),
3. **The role of a companion** (tool to assistant, companion, partner),
4. **Social skills** (not required or desirable to essential and necessary).

Two different paradigms can be distinguished regarding the potential relationships that might arise between humans and robots: the **caretaker paradigm** and the **assistant/companion paradigm** (Dautenhahn 2007, p. 698). The former states that humans take on a caretaking function when they encounter robots, since the latter are viewed as artificial creatures. In other words, humans look after robots (and not vice versa, like *care robots*). According to this robot-centred view, the human must recognize the robot's needs and react to them. The assistant/companion paradigm focuses on the robot's supportive role as a useful machine that is able to recognize and respond to human needs to be helpful. Thus, an assistant robot that supports people with everyday tasks takes on a role similar to that of a personal guardian or butler. The choice of which paradigm to emphasize in the design of a human–robot interaction is ultimately left to the developers of technology. The *poliTE* project and the **FASA MODEL**, however, take the view of the assistant/companion paradigm and focus on the human-centred perspective.

As an approach to concrete applications in specific interaction situations and application contexts relating to socially intervening artificial assistants, the **applications of the FASA MODEL are acted out below with examples** in the context of various settings that reflect interactions **between humans and artificial social assistants**. To do this, a series of case studies identified over the course of the literature search are presented. At the end of the chapter, we will consider a fictitious application example, or thought experiment, to illustrate the application of our model to the analysis of a concretely planned interaction situation (for example for a research project) between a socially intervening artificial assistant and a person. By that, we will show how to use the model as a checklist or for a research and development project.

6.1 Application example 1 – A robot barista

Hedaoo, S., Williams, A., Wadgaonkar, C., & Knight, H. (2019). A robot barista comments on its clients: social attitudes toward robot data use. In Proceedings of the 14th ACM/IEEE International Conference on Human–Robot Interaction (HRI), pp. 66–74.

In this application example, a robot barista (NAO, Softbank Robotics) commented on a conversation between two guests in a café setting. The valence of the robot’s comments varied (positive/negative). The addressees of the comments were also varied, as well as the setting of the conversation, e.g., family setting vs. job interview. The basic setting of this application example constitutes a framework within which the test subjects judge the appropriateness of the robot’s behaviour; therefore, some of the factors and factor criteria of the *FASA MODEL* of social appropriateness are already reflected: the robot is already configured for a specific situation with specific relationships between interacting agents and with individual specifics. It is intended to act:

- as a **barista for guests** (factor criterion of the «Relations between Interacting Agents» factor: <familiarity or relationship aspects>) in a café setting (factor criteria of the «Situational Context» factor: <place>, <framing>, <participants>)
- as a **conversation partner** (factor criterion of the «Individual Specifics» factor: <individual shaping of social roles>).

The appropriateness judgements of the test subjects also illustrate the relevance of the factors and factor criteria for human–machine interactions: the judged appropriateness of the robot’s behaviour depended on the social roles and degrees of

familiarity with which the test subjects met. The factors «Individual Specifics» and «Relations between Interacting Agents» therefore played a particularly influential role in judging the appropriateness of the robot's behaviour. In addition, the perception of the appropriateness of the robot's comments fluctuated depending on the subjects' mood in the context of the specific test situation – another link showing the relevance of the «Individual Specifics» factor. The fact that the valence of the robot's comments affected judgements of appropriateness also points to the «Type of Action, Conduct, Behaviour, or Task» factor: some (speech) actions have typical consequences that always manifest, except in unusual situations or unusual usage. Praise and niceties are usually met with a positive reception. The «Standards of Customary Practice» factor also plays a role in this application example: the test subjects' judgements of the appropriateness of body language, as well as potential conversation analysis and database queries done by the robot reflect a) the subjects' own implicit conceptions of appropriateness, which, if explicated, would presumably allow conclusions about specific customs depending on their social position, habitus, etc., and b) their personal conception of intimacy and privacy preferences. The aspects are related to known legal and ethical questions in the context of human–machine interactions and, thus, also demonstrate that questions of legality, ethics, and social appropriateness are in some cases closely related, although we should not be too eager to conflate them for this very reason.

6.2 Application example 2 – Baby schema

Mussakhojayeva, S., Zhanbyrtayev, M., Agzhanov, Y., & Sandygulova, A. (2016). Who should robots adapt to within a multi-party interaction in a public space?. In Proceedings of the 11th ACM/IEEE International Conference on Human–Robot Interaction (HRI) (pp. 483–484).

In this application example, people in mixed groups (parents, their children, and people with no relation to the families) met a robot (NAO) in field tests. The robot adapted its behaviour to either the children or the adults in the group as it presented itself. The results of the experiment showed that the robot's behavioural adaptations were evaluated differently by different people. Regardless of its verbalizations, children perceived the robot positively, whereas the parents perceived it more positively if it adapted its behaviour and language to the children. In general, the robot was perceived more positively by the parents than by the unrelated participants. This difference was very likely rooted in the children's positive reaction to the robot. The unrelated adults (without children) remarked that the robot

should adapt to adults in settings such as banks and hospitals, whereas the parents preferred the robot to adapt to their children, regardless of context.

This case study clearly illustrates the relevance of the «Type of Action, Conduct, Behaviour, or Task» factor with the «role identities» criterion since the relationship between the parents and the children led to differences in judging the robot's behaviour. Finally, the relevance of the «Situational Context» factor (situational criteria «space» and «framing») can also be seen since people with no relation to the children present in the interaction preferred behaviour adapted to adult interaction partners in 'official' settings such as banks or hospitals. However, this factor appears to play a subordinate role here because the role identity of being a parent shifted this preference in favour of unqualified adaptation to the children.

6.3 Application example 3 – That robot touch

Hoffmann, L. (2017). That robot touch that means so much: On the psychological effects of human–robot touch (Doctoral dissertation, University of Duisburg-Essen, Germany).

In our third application example, the influence of different parameters on the perception of contact between humans and robots and the effect of this contact on the evaluation and perception of the robots was tested. The results showed that touching certain parts of the body, e.g., the back or the legs, was perceived to be more appropriate from a robot than from a person (stranger). Furthermore, touch initiated by the human was perceived to be more appropriate than reciprocal touch or touch initiated by the robot. The acceptability of touch varied with the size and mechanical appearance of the robot; the appropriateness decreased as the size and mechanical appearance of the robot increased. In addition, touching the robot generally led to positive affect and more positive interaction behaviour. Accordingly, touch appears to be positive in HRI, but any touch should be initiated by humans. For example, it would be conceivable to establish a human-initiated handshake as a greeting.

This example is especially relevant as it demonstrates that a judgement of social appropriateness can differ between humans and robots. Here, a touch that would be inappropriate for a human is judged to be appropriate in the interaction with a robot. In terms of the FASA model, the «relations between interacting agents» factor is again especially relevant in this case study, with emphasis on the «familiarity and relationship aspects» criterion, which appears to be viewed more 'narrowly' with humans than with robots in the case of physical contact. Here, this factor goes hand in hand with the «Standards of Customary Practice» factor and

more specifically its <values/social norms> criterion. Simply touching a stranger violates socially established norms of behaviour. This is especially true for certain parts of the body, namely the back and the legs in this case study. However, the extent of the applicability or inapplicability of these norms seems to differ between robots and humans. The «Individual Specifics» factor also plays a role, given that the dependency of social appropriateness on the size and mechanical appearance of the robot was, in turn, contingent on personal evaluation structures.

6.4 Application example 4 – If a robot comes down the hallway...

Lauckner, M., Kobiela, F., & Manzey, D. (2014). 'Hey robot, please step back!'-exploration of a spatial threshold of comfort for human-mechanoid spatial interaction in a hallway scenario. In Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication, (pp. 780–787).

Our fourth application example examined the distance perceived to be appropriate when passing someone in a limited space (in this case, a hallway). The robot used was a prototype of a Bosch transporter robot, equipped with a display that could show a human face. It was found that the accepted proximity for a frontal approach was 0.8 m and the mean accepted distance for passing laterally was approx. 0.4 m. The preferred distance was not significantly influenced by the robot's autonomy but increasing the robot's speed (by 0.8 m/s) increased the preferred distance significantly. The robot's external appearance had no significant influence on the frontal distance, but a human-like design reduced the preferred lateral distance by 0.1 m. Due to interindividual variability, a frontal distance of 1.1 m and a lateral distance of 0.6 m was recommended for first contact with a social robot in this example.

In terms of the FASA model, the «Situational Context» plays an especially prominent role here (where something is unfolding, how it is spatially arranged as it unfolds); the frontal distance perceived to be appropriate was roughly twice as large as the accepted distance for passing laterally. The «Individual Specifics» factor and the <personal evaluation structures> criterion are also important, as they encompass differences in preferred distances from individual to individual. Additionally, the «Standards of Customary Practice» factor can be cited, since the <habitus> criterion reflects the field of proxemics, which was the focus of the study. Thus, there are certain distances that have been socially established as typical (Hall 1966), and failing to observe them constitutes a breach of social appropriateness.

6.5 Application example 5 – Wait for it... Hello!

Yamamoto, M., & Watanabe, T. (2006, September). Time lag effects of utterance to communicative actions on CG character-human greeting interaction. In Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication, (pp. 629–634).

In our fifth application example, a series of differently timed variations of a greeting were examined in a Japanese-speaking setting. The robot was small and somewhat playful (*unazuki-kun*) and represented the embodiment of a virtual agent. During communication, it was found that variations in pauses and delays led to different communication effects. Thus, a delay of 0.3 s was desirable for greetings between acquaintances, but longer delays were preferred for polite greetings.

In terms of the *FASA MODEL*, the «Situational Context» factor (situational criterion «place») is reflected in the Japanese cultural setting. There are greater differences between polite and familiar greetings in Japanese culture than in some other cultures (e.g., the American culture), but the preferred delays would most likely manifest in some other way in other cultures, or perhaps play no role at all. The «familiarity and relationship aspects» criterion of the «Relations between Interacting Agents» factor is also relevant, reflecting the social distance between the two interaction partners. The preferred delay in communication was defined by the familiarity of both persons (acquaintances vs. interacting agents less familiar with one another), which in turn determined whether a formal, polite greeting was viewed as socially appropriate. The «Standards of Customary Practice» factor, more specifically the «habitus» criterion, can also be used to understand social appropriateness in this application example. The «habitus» criterion encompasses the typical and ‘ingrained’ types of behaviour and judgement within a group, which includes the general rules of conversation.

6.6 Application example 6 – CLIPPY

Whitworth, B. (2005). Polite computing. Behaviour & Information Technology, 24(5), 353–363.

The final application example relates specifically to ‘Clippy’, Microsoft’s virtual assistant. Clippy was perceived as impolite and disruptive by many users due to frequent and unsolicited interruptions in the work process, despite its potentially positive function in terms of assistance. This perception of it being impolite and disruptive may stem from it not being designed with enough thought for

interaction performance with the user. Whitworth writes: “Politeness is any unrequired support for situating the locus of choice control of a social interaction with another party to it, given that control is desired, rightful and optional” (Whitworth 2005, p. 355) and formulates the following informal politeness rules for software (which can largely be transposed to robots) (cf. Fig. 6.1):

According to Whitworth, Clippy violated a number of these postulated informal politeness rules while acting as an assistant and was consequently perceived as disruptive and annoying by many users.

In terms of the *FASA MODEL*, the «Situational Context» factor and its situation criterion «media-based and performative mediation» are especially relevant here since the norms of e-conversations fall under this criterion. For virtual or media-based conversations, there are specific norms that must be considered when designing interactions within this context to ensure that they are perceived as socially appropriate. The «Standards of Customary Practice» factor and its «habitus» criterion, which encompass the typical and ‘ingrained’ types of behaviour and judgement within a group, including general rules of conversation, are also relevant. The interaction between the different factors is clearly visible in this example. Although the «Standards of Customary Practice» factor includes general rules of conversation, the situation, in this case the media-based interaction, means that another set of norms, some of which are different and some of which overlap with the usual norms, must be considered.

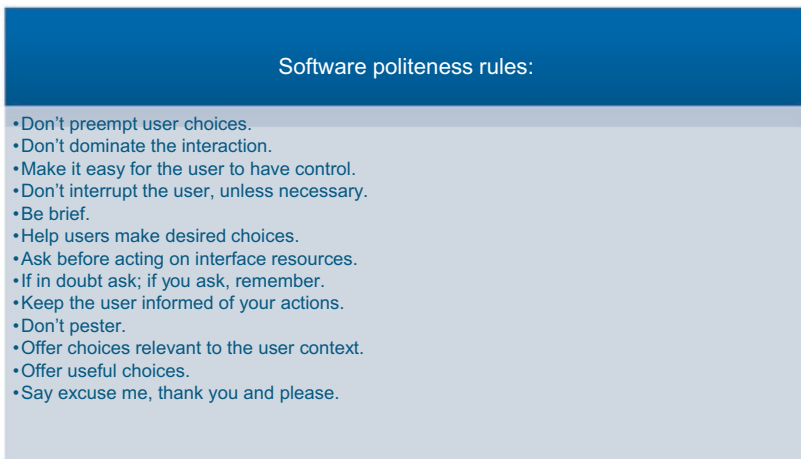


Fig. 6.1 Software politeness rules (Whitworth 2005, p. 359.)

6.7 Fictitious application example/thought experiment

To illustrate the **application of the model** to planning specific interaction situations in the context of **research and development**, it might be helpful to consider a fictitious example as a thought experiment. The FASA model can be used as a basis to assess which aspects of social appropriateness should be taken into consideration. We shall examine the criteria factor by factor to determine whether each criterion is relevant using the questions listed in the model description and, if so, what consequences we can deduce in terms of the behaviour that we wish the system to perform. As a realistic and widely encountered scenario in state-of-the-art technological development, we will consider the implementation of a **robot in a retirement home**. We will apply the model to a concrete interaction situation: reminding the residents about an appointment, in this case an upcoming leisure activity that has been planned.¹ The following discussion represents *one* possible approach to this scenario based on the *FASA MODEL* and makes no claims of exclusivity or completeness.

«**Situational Context**»: For our appointment reminder, let us begin with the factor «Situational Context» and analyse the situation using the factor criteria listed in the model.

«**Place**»: Our analysis begins with the situational criterion «place», which relates to where the behaviour is taking place. Here, the specific context of the retirement home and the cultural setting in which it is located must be considered. In our example, we will assume the case of a Western culture. If the interaction is set within a German-speaking country, the polite *Sie* form of address should for example be used if the level of familiarity is low (see also «Relations between Interacting Agents»), and a larger interindividual distance should be maintained than in some other societies (see also the «habitus» customariness criterion, which encompasses proxemics). Given that the interaction is set within a retirement home, we can draw conclusions about the age of the interacting agents and

¹ We must bear in mind that, especially in care work, ethical aspects play a major role when designing behaviour and must be considered accordingly. This thought experiment remains deliberately vague in some regards to avoid any assumptions that could potentially be interpreted as ableism or ageism. The extent to which any physical, cognitive, or other restrictions should be considered in behavioural design is specific to each situation and must be judged by care workers on a case-by-case basis.

the possibility that the addressees may have cognitive and physical limitations caused by degenerative diseases or simply advanced age (see also «Individual Specifics») that must be taken into consideration when designing the interaction. Furthermore, within the scope of the «place» situational criterion, the degree of publicness (e.g., private vs. public) of the interaction needs to be considered; in the application example, the situation can be assumed to be private, unless the human-robot interaction is unfolding and being recorded as part of a scientific study. Accordingly, the robot does not need to communicate as representatively as would be necessary in a completely public situation.

«Framing»: The situational criterion «framing» asks ‘as what’ the behaviour is unfolding. For example, this criterion encompasses whether the action being evaluated is a ritual or ritualized, and whether it is being performed for its own sake or as a form of commentary, e.g., in the context of art. In our application example, the level of seriousness is most important aspect to consider. The situation is not very formal, and we can assume that it will occur somewhat regularly. Therefore, the robot’s behaviour can be more informal, jokes or possibly colloquial language are permissible, and no lengthy explanations are required, since the addressees can be assumed to be familiar with the situation.

«Media-based and performative mediation»: The situational criterion «media-based and performative mediation» plays a subordinate role in this situation since the interaction is unfolding in a face-to-face setting. This criterion would play a more prominent role in media-based interactions such as video conferences or discussions in a comments section on the internet.

«Participants»: For the «participants» factor criterion, which encompasses the nature of the interacting agents, we simply need to consider that the participants are people and no entities of a different nature play a role or need to be considered.

«Time»: The situational criterion «time» describes whether the appropriateness of a displayed behaviour depends on it being performed at specific times. This criterion also plays a subordinate role in this example, since the performance of the robot’s reminder task does not depend on the specific time at which this reminder is given (unless the reminder would undesirably wake the residents from sleeping, etc.). It is sufficient to select a time window that allows the addressees to complete or interrupt their current activities to participate in the planned leisure activity, or travel to the necessary location.

«Individual Specifics»: As mentioned above, the «Individual Specifics» factor plays a prominent role in this application example.

«Personal evaluation structures»: The «personal evaluation structures» criterion focuses on aspects that influence whether an interaction partner would judge a behaviour as socially appropriate or inappropriate. In the context described here the age of the addressees must be considered, since the conditions of socialization differ from generation to generation. For example, a higher degree of formality in the form of address may be advisable when interacting with older persons, whereas people of younger age or from a later generation might perceive a more informal address (such as the German *Du* form of address) as appropriate, even in the absence of familiarity (see «Relations between Interacting Agents»). Physical and cognitive states also need to be considered, for example in relation to the speaking volume and speed that would be considered appropriate. A hearing-impaired person would judge a higher volume to be appropriate; in the context of a retirement home, a clearer speech style or higher volume might be considered appropriate, depending on the composition of the group of residents – or come across as discriminatory. Furthermore, regarding physical condition, a resident with restricted mobility might need more time to cover distances, which should be taken into consideration when defining the timing of the reminder for the specific group being addressed.

«Personal characteristics»: The factor criterion «personal characteristics» describes the dispositions that interaction partners bring with them. This includes personal preferences regarding certain aspects of the interaction situation. Likes, prejudices, and personal taste play a role, as do chronic distortions of perspective (e.g., a *negativity bias*), personal attitudes, understanding of irony, personal interests, etc. Since our example concerns a group interaction where the robot addresses multiple people at the same time, this criterion plays a less prominent role. Instead, it makes sense to design the interaction more generally, as it is unrealistic to expect to be able to account for the personal characteristics and preferences of all addressees. Nevertheless, it would be conceivable to have a scenario in which a particular person requires special attention to encourage them to participate in the planned activity. A two-stage process could then be envisaged that first makes a general address to the full group, then addresses a particular individual with a more personalized communication that considers their personal preferences to improve the perception of appropriateness.

«Individual shaping of social roles»: The criterion «individual shaping of social roles» is not too relevant in our example, since potential social roles are primarily meaningful when interacting in a context where the interacting agents have essential roles that determine the interaction itself. Conceivable examples include interactions within professional life, where there are superiors and employees,

interactions in school contexts, where there are teachers and students, and so on. In the interaction situation described here, which takes place in the context of a retirement home, social roles have a less prominent meaning because the addressees are unlikely to have any other roles besides being a resident of the facility at this point in time.

«**Type of Action, Conduct, Behaviour, or Task**»: Next, we need to analyse the «Type of Action, Conduct, Behaviour, or Task» factor and its factor criteria.

«**Time**»: Let us begin with the action and behaviour criterion «time», which relates to how and when the behaviour is taking place. Various questions could be relevant here. For example, what type of conversation is it? A dialogue or something else? Is there a fixed exchange between the interacting agents? Are there interruptions? If so, why? In some cases, this criterion could also encompass a sense of tact regarding choosing the right time to address a particular question within the interaction. In the application example, the most relevant aspects are: it is not unfolding as a classical dialogue with a fixed exchange between the members of the conversation, and it is not an instance of dyadic communication, but a communication with a group of people. In such a context, since the robot cannot necessarily wait until nobody is speaking to avoid interrupting a conversation, it could be viewed as socially appropriate to interrupt existing interactions with an ‘interjection’ before ‘broadcasting’ a general announcement of the upcoming appointment to the room – like how it is appropriate to tap on a glass to interrupt conversation before giving a speech. Furthermore, a sense of tact is not necessary for the robot since the subject of the interaction is not sensitive.

«**Role identities**»: The second action and behaviour criterion, «role identities», is closely linked to the «Relations between Interacting Agents» factor, as well as the individual criterion «shaping of social roles». This criterion relates to questions about who is performing the behaviour and who is judging it. Here, for example, it corresponds to aspects regarding the innate roles of the interacting agents or the roles assigned to them, how a judgement of appropriateness depends on the role of the person performing the relevant behaviour (e.g., depending on age, gender, or ethnicity), how the role identities of the interacting agents relate to one another, the reputations of the interacting agents, and whether people are representing themselves or somebody else within the interaction (this plays an especially important role in political settings). However, since social roles are not that important in this example, as mentioned earlier, this factor is irrelevant or only plays a very subordinate role.

«Intention»: The action and behaviour criterion «intention» relates to the motivations, goals, or intentions with which the behaviour or action is taking place. Here, it should for example be considered whether the behaviour has a persuasive intention, whether there is the potential for cooperation, etc. In the example, the robot could potentially also have a persuasive intention in addition to its reminder function to motivate the residents to participate in the planned activity. In a human context, a persuasive intention can be pursued in subtle and socially appropriate ways by strategies such as particular rhetoric or by mirroring gestures or facial expressions in communications, but that is not yet possible to the same for robots. Since the ability to recognize emotions and situations is also significant in this criterion, besides the ability to cooperate, its relevance may increase as technical development continues to progress.

«Consequences»: Finally, there is the «consequences» criterion of the «Type of Action, Conduct, Behaviour, or Task» factor, which encompasses the possible consequences of the interaction. Here, rules of conduct established within specific groups of actors can be considered, as well as the visibility of consequences or the (institutionally normative) enshrinement of behavioural rules (see also the «Standards of Customary Practice» factor). The power dynamics between the interacting agents also play a role; they determine whether a violation of the rules of social appropriateness can be sanctioned by the interaction partner, and, if so, how. The worst-case scenario in the example considered here would be a termination of the interaction, for example if the residents simply ignore the robot and do not participate in the planned activity. Accordingly, the consequences are not serious and there is some leeway for social (in)appropriateness in the design of the interaction, which would not be the case for other scenarios that carry more serious consequences (imagine for example a situation unfolding in front of a court).

«Relations between Interacting Agents»: The «Relations between Interacting Agents» factor also contains various factor criteria that must be considered when designing the interaction.

«Familiarity and relationship aspects»: The relational criterion «familiarity and relationship aspects» asks how the interacting agents relate to one another. This includes consideration of the frequency and duration of the interaction, whether a friendship exists or how close the interacting agents are to one another, what specific power dynamics there are, what expectations they have of one another, etc. In the application example, the robot and the residents already know and have frequently seen one another, and interaction situations have already occurred repeatedly. This allows us to conclude that no highly formal communication or

introductory greeting and self-introduction are needed. Since the example does not involve a *companion robot*, and the robot is instead understood as a service provider, the interaction does not have any special intimacy or familiarity and should instead be kept relatively neutral. There is not expected to be any power imbalance between the interacting agents; this aspect can therefore be neglected. Although the robot gives a reminder of the appointment, its task is not to force the residents to participate, nor does it have the authority to do so; it is simply offering a suggestion.

◁Intention: This is also reflected in the ◁intention> relational criterion, which describes the interests at stake in the interaction. For example, this includes the question of whether the interaction is cooperative or competitive, whether it serves an economic interest, whether the possible consequences of the interaction are institutionally enshrined, etc. In the application example, no further aspects need to be considered.

◁Context: The relational criterion ◁context> is comparable to the «Situational Context» factor but focuses instead on the relations between the interacting agents. For example, ‘as what’ do the interacting agents perceive the interaction? Is there a consensus on this perception? Applied to the situation of a robot reminder, it should for example be noted that both parties need to perceive the reminder as voluntary, i.e., the fact that the leisure activity is being proposed as a suggestion needs to be communicated to maintain social appropriateness.

◁Standards of Customary Practice: Finally, we must analyse the «Standards of Customary Practice» factor.

◁Values/social norms: The first customariness criterion is the ◁values/social norms> according to which the social appropriateness is judged. This criterion considers more collective values and virtues, such as fairness and equality, as well as individual values that the interacting agents may have internalized over the course of their socialization, but which are not necessarily shared by others, and finally any institutionally enshrined customs. In the context of our application example, this factor is of moderate importance, since the situation is not particularly sensitive to any potentially applicable values, but the violation of such a value might still lead to the interaction being terminated. For example, the fact that values such as friendliness or politeness can play an essential role in a retirement home setting needs to be considered. In addition, in the application example, no separately institutionalized customary practices are expected, although any relevant legislation should of course be observed, especially as it relates to the provision of care.

◀Habitus▶: The next criterion is the ◀habitus▶, which describes the types of behaviour and judgement ‘ingrained’ in a group. This is linked to the frequency of interaction. People who are meeting for the first time (excluding broader societal contexts resulting from perceived roles in the interaction situation) do not have any habitually ingrained types of behaviour and judgement. In our application example, it might have already been established as common practice for the leisure activities to begin slightly late, for instance, so it would accordingly not be socially inappropriate for the robot to account for this time window in its reminder. This criterion also encompasses types of behaviour that would otherwise be inappropriate but have been accepted as appropriate through the habitus of the relevant group (but would not be accepted in other group contexts).

◀Regulative norms▶: The final customariness criterion ◀regulative norms▶ relates to the ethical dimension of the application example. The example could potentially involve vulnerable groups of people, making this criterion especially important. A long-standing and independent area of applied ethics exists for this scenario, identified by the keyword of ‘care ethics’, so we can draw from both the existing professional discourse and the relevant experts and institutions on this topic.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Reflections on Sociosensitive and Socioactive Technology

7

Chapter 2 gave a list of purposes for which this book and the FASA model are suitable and another list for which they are not. But regarding the model's applicability and validity and in general, there are various other questions about sociosensitive and socioactive technology that must be considered, including ethical aspects. This book also aims to increase awareness of the complexity and sometimes problematic nature of developing, designing, disseminating, using, and regulating sociosensitive and socioactive systems. The chapter at hand therefore presents some initial reflections about sociosensitive and socioactive technology. It does not make any claim of exclusivity or completeness; on the contrary, it should be understood as an appeal for non-technical expertise to be continuously and more thoroughly integrated into technology design projects. These reflections bring the book to a close and serve as a prelude to subsequent discourse.

7.1 Social appropriateness and tact

When saving face is a concern in social interactions, one component of socially appropriate behaviour is what is known as 'a sense of tact'. Regarding technical systems, we might ask whether it makes any sense to say that a technical system is capable of 'embarrassing' us, and whether experiencing embarrassment towards a technical system can itself be appropriate. Helmuth Plessner defines tact as

the ability to perceive imponderable differences, the ability to grasp that untranslatable language of phenomena spoken by people without words through their constellation, their behaviour, their physiognomy in the unfathomable symbols of life. Tact is the willingness to respond to the finest vibrations of the environment, a willing

openness to see others and thus take yourself out of focus, to measure others by their own standards and not your own [cf. «Individual Specifics», remark by the authors]. Tact is the eternally watchful respect for the souls of others and hence the first and last virtue of the human heart. (Plessner 2002, p. 107, own translation)

It will certainly be difficult to endow technical systems with such a watchful respect for the souls of others, and so we must ask ourselves whether and when it makes sense to attempt to simulate such respect. Here, it is useful to make a distinction between two different forms of social appropriateness, both of which we have discussed in greater depth elsewhere (Bellon et al. 2022). Socially appropriate behaviour can relate to respect for the dignity of the other in a very existential sense. But it can also refer to a weaker, possibly derived, form that facilitates interactions and society through conventions that are constantly being renegotiated (cf. for this and Niklas Luhmann's notion of tact Youssef 2019).

According to Hans-Georg Gadamer, tact can be understood as a kind of '**social sensitivity**' (Gadamer 2011, pp. 13–15). The following excerpt by David Kaplan insightfully reveals the connection between tact/politeness, education, self-cultivation, and appropriate technique. Some of the factors of the FASA model are also reflected in the excerpt:

What is this sense of appropriateness? For Gadamer it is "tact." It is a particular kind of social sensitivity to social situations and the judgment of how to behave in them. Tact is the tacit knowledge of appropriate action for a particular circumstance. It involves knowing what to say and do – and what not to say and do. Although not based on general principles or universal concepts, Gadamer maintains tact is a universal sense that requires of all that we remain both sensitive to particular situations, guided by the wisdom of the past [handed down as customs, remark by the authors], yet open to other points of view. Although it is difficult to prove any matter of tact conclusively, it is not an irrational concept; it is merely an acquired ability. How does one acquire it? Through education in culture, development, and self-cultivation in society. That is to say, *Bildung*. The only way to acquire interpretive tact is through practise. This connection between tact and practical wisdom has completely dropped out of the contemporary conversation of technology. But what is largely at issue in questions concerning the good life in a technological age is this notion of appropriateness in conduct. Technology is shot through with tact. It answers key questions, such as how things ought to be designed, how they should be used, how they should affect others, how they should be governed. Tact may not provide a precise answer to any of these questions, but if universalist and scientific concepts are ruled out (or not exclusively employed) then all that is left is practical wisdom, developed over time, through *Bildung*. After Gadamer, the notion of 'appropriate technology' takes on a whole new dimension. New answers might be found to vexing practical questions concerning technology. (Kaplan 2011, p. 232)

The **five factors proposed in this book** and their respective criteria not only allow an understanding of Gadamer’s concept of tact, **but, as the fundamental factors of social appropriateness, they also offer an approach to the question of how behaviour is judged to be socially appropriate or inappropriate in interpersonal and human–machine interactions.** First, as proposed here, the behaviour in question must be perceived through observable aspects such as voice and tone, gestures and facial expressions, posture and positioning in space, and so on, all of which vary in time, space, and mode. Which manifestations of these observables can now be understood **as** indicators for appropriate or inappropriate behaviour primarily depends on the five factors of the model: «Type of Action, Conduct, Behaviour, or Task», «Situational Context», «Relations between Interacting Agents», «Individual Specifics», «Standards of Customary Practice». If this model and the research results of the *poliTE* project summarized in this book are adopted by future concrete research, especially in technology design projects, then – it is hoped by the authors of this book – various typical and meaningful combinations of observables, indicators, factor criteria, and factors could emerge. Besides being useful for the design of sociosensitive and socioactive systems, this could also provide further insight into the phenomenon of socially appropriate behaviour, appropriateness cultures, and the normative foundations and conditions of plural life forms in digitized life realities.

7.2 Why social appropriateness in human–machine interactions?

If insight into the connection between judgements of appropriateness and significant groups of observables typically associated with specific contexts, tasks, and social relationships is possible in this sense, how can this contribute to the design of socially compatible technology? Could self-learning sociosensitive or socioactive systems even record and compile these relationships, allowing the factors of appropriate behaviour to in turn be researched through an analysis of the information compiled by systems? Could these observables be made machine-readable, i.e., made to be understood by systems as indicators – if so, which of them? Enough of them? Could processing at the level of indicators, i.e., **reading observables ‘as symbols for’** and therefore the interpretive steps that connect an observable with a judgement of appropriateness, be partially or even mostly delegated to highly automated systems, allowing this ability to be integrated into sociosensitive or socioactive systems? And if it is indeed possible to have systems with

the ability to establish judgements about the social appropriateness of actions and behaviours, then we must reflect upon the technology by asking ourselves the question:

Should such systems be developed and used in the first place?

This book compiles some examples of observables that are currently being considered in research and prototype implementations of interactive systems, such as ‘system judgements’ about the appropriate distance between the participants in the interaction (proxemics). All research efforts into emotion-sensitive adaptive systems, human-like interactions, artificial assistants, and *companion technology*, social robotics, etc. share one – more or less explicit – premise:

Human-like or ‘natural’ – or at least less artificial – interaction is better. Is this true?

Why would someone build systems that simulate human behaviour up to or beyond the *uncanny valley* (Mori 2012)?

7.3 Sociosensitive/socioactive systems as seemingly human?

From the perspective of technology assessment, such sociosensitive and socioactive systems clearly present challenges, such as the possibility of deception (is something a person or just technology?) that leaves users unsure about the true nature of the entities with which they are interacting. Such systems would pass the Turing test, at least in the short term. Although this ambiguity may lead to conscious and enjoyable immersion in human–machine relationships, at the same time, fundamental respect for humans as well as human rights (and labor rights and conditions) can be jeopardized wherever AI systems are enacted and assumed in roles of humans and instead of humans. This is strategically exploited by many data-driven or platform-based services, some of which employ human workers in precarious conditions while presenting themselves as purely AI-based functions to users. This has been described as **ghost work** by a new global underclass (Gray and Suri 2019) or as a deliberately staged impression of magic that makes work seem like it is ‘plucked from the cloud’ and performed by ‘magical hands’.

The Amazon version [of Mechanical Turk, remark by the authors] is a way to easily outsource – to real humans – those cloud-based tasks that algorithms still can't do, but in a framework that allows you to think of the people as software components. The interface doesn't hide the existence of the people, but it still does try to create a sense of magic, as if you can just pluck results out of the cloud at an incredibly low cost. (Lanier 2013, p. 169 f.)

This impression of magic could undoubtedly be considerably enhanced using socially appropriate artificial agents, causing the precarious employment conditions affected by it across the globe to fade even further from the sight of potential users and consumers. Though users and customers may find such moral convenience attractive (it is not pleasant to see the misery 'behind' products), efforts to raise public awareness might be necessary – the social media platform *Facebook* employs people and not algorithms to weed out pictures showing decapitation and torture, sometimes at a considerable psychological price (The Verge 2019) – and a sharper public understanding of sociosensitive and socioactive technology might also be needed.

Another consequence of sociosensitive and socioactive agents – corresponding to either an opportunity or a risk depending on the realizations of the factors – is the facilitation and **increase of parasocial relationships**; this term describes relationships where people fall in love with fictional characters like James Bond or non-human entities like God or robots. Viewed as an opportunity – consider for example relationship or sex robots – this field represents a billion-dollar market that has the potential to alleviate loneliness. But parasocial relationships can also have problematic consequences, for example if a beloved robot (or car) is 'rescued' instead of another person, or 'material damage' to a beloved robot partner is met with pre-emptive or retaliatory bodily harm – perceived as self-defence. There are of course non-problematic ways to develop attachment to objects beyond technical systems; nevertheless, the design of technical systems should consider any potentially problematic consequences that can be anticipated.

Another challenge of minimizing human-technology differences in human-machine interactions with sociosensitive or socioactive agents is not to overemphasize this minimization of differences. For human-machine interactions, which involve learning systems on the technological side and have long since taken place on the basis of 'comprehensively networked IT systems' (Wiegerling 2016)–and even for implementations that exploit big data technologies, which are

sometimes touted as almost magical—the following principle still holds true (Richter und Kaminski 2016): when interacting with a technical system, people are only **addressed as tokens of various compiled profile types** (ranging from hard-coded default users to average users with various adapted characteristics to fully ‘personalized’ users) by the system.

But reducing human–machine differences in human–machine interactions with sociosensitive and socioactive technology could obscure this typified default type of human–machine interaction, leading technical systems to be classified as full social actors. This may already seem like an attractive option in socially sensitive areas, not least due to economic incentives. But such a misapprehension of sociosensitive and socioactive technology as full social actors is undoubtedly fraught with risk (Nähr-Wagener 2020): suppose that a person is unable to articulate individual wishes and feelings in areas where articulating one’s own mental state is essential and should ideally be facilitated¹ (e.g., in nursing and care work) precisely *because* the sociosensitive and socioactive system is unable to process them adequately: talking to a system about ones ‘emotional inner life’ is pointless, and so a person might eventually just stop trying to articulate this desire. In this scenario, the person may at some point no longer consider it worthwhile to develop personal wishes and feelings in the presence of permanent human–machine interactions of this type – there is a danger of self-reification. Even sociosensitive or socioactive care robots cannot be empathetic or sympathetic interlocutors. If we only consider the point of view described above, assistance from sociosensitive or socioactive care robots should presumably be restricted to only classical assistance tasks, in particular bureaucratic tasks. Thus, sociosensitive and potentially socioactive care robots should also act as assistants for care workers, rather than as independent care systems that might even participate in relational work. In general, this means: the boundaries of possible fields of application of sociosensitive and socioactive technical systems should not be determined by the illusion that these technical systems constitute social actors in a comprehensive sense.

¹This remark is intended in the following sense: in practice, especially in the nursing sector, patients often do not have enough time to articulate their own mental states (‘nursing crisis’), meaning that trends of self-reification are undoubtedly already a reality, even without sociosensitive and socioactive care robots.

7.4 Advantages of sociosensitive/socioactive technology/ Why and why not?

Systems that take social appropriateness into consideration hold considerable potential for **more pleasant human–machine interactions**, which in turn can have positive effects on people’s mood, health, motivation, performance, etc. The aforementioned research on respect (Sect. 4.4.1.1), for example, considers a recipient-based concept of respect according to which respect exists when a person **feels respected**, regardless of whether the relevant interacting party truly respects them (Quaquebeke and Eckloff 2010).

In a fundamental sense, technical systems can never respect someone.

For example, they are not in a position to choose who to respect and who not to respect, because they have no authentic normative preferences to guide such a choice. Such systems could be implemented with a corresponding axiology from which attestations of respect could be derived, but the processing of foreign values implemented by foreign rules would not represent respect in the sense of **autonomous recognition**, which presupposes a conscious subject capable of recognition (cf. for example Siep 2022; Gransche et al. 2014). But although **technical systems cannot genuinely respect people**, they could potentially **simulate expressions of respect sufficiently well** that a person feels respected by an artificial agent. According to the recipient-based concept of respect, this would be sufficient to induce the positive effects of respect on health, motivation, performance, etc. in the person who feels respected wherever this is desirable.

Another possible advantage of sociosensitive or socioactive technology that provide services is that inappropriate conversation interruptions by interactive robots or AI systems in hybrid social settings, e.g., conference coffee breaks served by robots, could be reduced. This would not necessarily or not just improve the quality of human–machine interactions but more importantly the quality of interpersonal interactions. Such an advantage would however only prove fruitful if interactions between humans and technology in this kind of hybrid setting is desirable or necessary for other reasons (e.g., economic), since technology with improved sociosensitivity or socioactivity only mitigates technical imposition that would not be disruptive in the first place if the technical agents (e.g., catering or care robots) were never introduced to the social context. Accordingly, **the promotion of technical sociosensitivity and socioactivity risks**

falling foul of a tech-fix ideology² (to which, in a self-critical sense, this book might also be contributing) where the purpose of version $n+1$ is merely to solve problems that were introduced by version n . We must never stop asking whether people are a better choice than attempted technological substitutes in the hotel reception, the service hotline, behind the counter, for practical psychoanalysis, in sports clubs, in intercultural training settings, as intimate partners, etc. One of the tasks of technology evaluation (in every sector: politics, science, economy, etc.) is to invest the finite resources of a society towards development goals that exist by consensus; for example, we must ask whether a precarious care system (or even a hypothetically perfect one) only needs care robots – if they are indeed needed at all – because precisely those preferences and images of society and humanity that facilitate the existence of sociosensitive and socioactive agents are the root causes of the crises plaguing the health system. In the research and development of interactive systems, the FASA model **can also be used as a heuristic** to decide in what «Situational Contexts», according to which (and in some cases overcoming which) «Standards of Customary Practice», for which «Relations between Interacting Agents», for what «Individual Specifics», and for which «Type of Action, Conduct, Behaviour, or Task» **sociosensitive and socioactive systems might in fact be worse** than socially indifferent systems, or indeed under what circumstances any technical system at all might be disadvantageous compared to solutions based on appropriately qualified people. In collectives with a tendency towards tech fix reflexes, i.e., a propensity to respond negative effects of technology with more rather than less technology, potential no-tech and low-tech solutions receive few resources, late resources, or no resources at all (research focus, development funds, etc.).

The factors and criteria of social appropriateness are so complex that adequate and conscious consideration should in many cases lead to the informed decision to refrain from any technical implementation of them.

² *Tech fix* is short for *technological/technical fix*, sometimes also called *solutionism*. It describes the attitude of primarily seeking to fix problems with technology, even when the problems were mostly caused by technology in the first place. For example, the problem of CO₂ or fine dust emissions could be tackled socially or at the level of people's actions by adopting practices that avoid generating emissions, instead of resorting to tech fix with improved particle filters or CO₂ separation systems. The problem with this, and the reason why tech fix or (technological) solutionism often carries negative connotations, is that technical solutions are often only necessary to solve technical problems that were originally intended as technical solutions for other technical problems, and so on and so forth.

7.5 Technology does not interpret and does not understand

In addition, only a fraction of the listed observables can currently be technically implemented and technically integrated as indicators for criteria. Although this may be improved in future for many observables through further research, it would make sense to clarify whether and which of these observables are suitable for being technically processed as indicators at all before investing in such research. As presented in Chapter 4, observables are not just detected, they must be **read as** indicators. Understanding something **as** something, i.e., understanding an observable manifestation *as* a symbol and *as* information about something unobservable, is an **act of interpretation**. If we wish for technical systems to assign a meaning to observables for social appropriateness beyond simply applying a fixed – and hence predefined – reference template (and the dynamic character and complexity of the phenomenon suggest that this is indeed desirable, see Bellon et al. 2022), then these systems must have interpretative abilities. Whether non-living entities can fundamentally bridge this **hermeneutic chasm of understanding**, even if they can simulate such abilities by processing, is an ongoing debate in the philosophy of technology (Gransche 2021; Romele et al. 2018; Romele 2020) and is at the very least questionable. Similar to the recipient-based perspective of respect, according to which respect that is felt without being truly given still has a positive impact, understanding (of something as something) might philosophically be impossible to achieve by technology, but the corresponding technical surrogates (e.g., simulated understanding) might still suffice to benefit from the positive effects of (simulated) socially appropriate behaviour by technical systems.

7.6 Politeness as blameless deception – *fake it until you make it*

Since politeness – as highlighted in the title of the *poliTE* research project – is an important element of social appropriateness, it is worth looking at the philosophical treatment of politeness as part of the decision of **whether and how technical systems should be made socially sensitive, or even actively polite**. Given the chasm of understanding mentioned above, critics might conclude that any plans to design sociosensitive or socioactive systems will ultimately prove in vain and only elaborate gimmicks or neo-baroque masterpieces of illusion like

the mechanical chess-playing automaton (“The Turk”) can possibly result from them (cf. Standage 2002). To such a general rejection of sociosensitive or socioactive technology, one could reply that in many cases (according to sufficiently many criteria), technical systems can compellingly display (simulate) socially appropriate behaviour despite the aforementioned chasm (even if this behaviour is not genuine). In other words, the variance in «Situational Contexts», «Relations between Interacting Agents», «Individual Specifics», «Type of Action, Conduct, Behaviour, or Task», and «Standards of Customary Practice» could be kept sufficiently low that the remaining complexity is at least provisionally represented through elaborate reference templates that could possibly also be updated through conditional learning.

So, what would be gained if technical systems regulated their interventions according to social appropriateness criteria, with the illusion of politeness – deceptive, but deceptively real?

A partial answer to this, at least in the area of human politeness, is offered by Immanuel Kant, who regards polite behaviour as a deception, but a ‘blameless deception’ (cf. Kant 2006, p. 43–44) that does not harm the ‘deceived’ party, since the deception is an open and **known cultural technique**. In the case of feigned respect that is still felt by the recipient, the deception would not only not be harmful, but it would even be beneficial in some circumstances (at least if the goal is improved efficiency at work, etc.). Kant offers a way in which politeness, feigned or simulated sociosensitivity or socioactivity, does not harm the actor while also benefiting the deceived; it “is nevertheless very beneficial as an illusion” (Kant 2006, p. 43). This approach could be summarized as: *fake it until you make it*.

In general, everything that is called *propriety (decorum)* is of this same sort – namely nothing but *beautiful illusion*. *Politeness (politesse)* is [...] to be sure not exactly always truthful [...] but this is precisely why they do not *deceive*, because everyone knows how they should be taken, and especially because these signs of benevolence and respect, though empty at first, gradually lead to real dispositions of this sort. (Kant 2006, p. 44)

For Kant, the simulation of appropriate behaviour (empty signs of benevolence) is a way towards skillfully appropriate behaviour or virtue (true conviction). Aristotle (2014, II, 6, 1106 b 36) already clearly showed that virtue is something that can be cultivated, i.e., developed through practise under the right conditions. One of the **basic conditions of our life is our embedding within and interactions with technology**. Every human–machine–world relationship has and has always had *deskilling*, *reskilling*, and *upskilling* effects, not only in the field of

professional skills – an area that has long been the subject of intensive research – but also regarding our basic judgement and moral skills (cf. Vallor 2015).

Moral skills can be understood as:

The ability to properly assess the proper behaviour towards the proper person and the proper time in the proper place in the proper way.

Being able to behave socially appropriately is a skill that also requires the ability of moral judgement. Analogously to the ability to interpret, we can ask whether moral skills are fundamentally inaccessible to technical systems – there are many reasons to think so that cannot be presented here. Besides the ability to act, normative judgement is essential to **not only be able to do something, but be able to do it appropriately, sensibly, and responsibly**. Moral skills are also learned rather than innate, they are developed by practising, provided that certain conditions are met, such as:

- the existence of role models,
- the opportunity for repetition,
- sufficient feedback,
- cognitive and emotional resources,
- motivation/interest.

If these conditions are met, competencies, (moral) skills, attitudes, and genuine virtue can be cultivated. However, precisely these conditions are threatened by modern technology in some regards. Learning systems, for example, deny the opportunity for repetition (since it is not possible to interact with the same system state twice), making it impossible to receive feedback that can be used for practising: if complex, networked, learning systems change their part in an action after every instance of interaction based on opaque control parameters (e.g., user behaviour, other users, the interests of the operator, environmental data, etc.), users cannot possibly learn to correct their own part in the action from the combined effect of the hybrid action. If someone cannot understand their own influence on a combined result, they cannot deduce the effects of changes in this influence and therefore cannot redirect their own influence to accomplish their goals more effectively (i.e., learning). Continuous exposure to increasingly powerful assistance services also leads to a gradual loss of ability in the delegated parts of actions (Gransche 2016): driving with a navigation system causes you to gradually forget how to navigate without one. Extensive use of low-threshold, communication-simplifying computer technology (for remote communications) carries the risk of losing the ability to engage in face-to-face communications

(possibly even pathologically), so that in the end, everyone might end up paradoxically “alone together” (Turkle 2011). Technologies influence the conditions of our potential to develop abilities, as well as the concrete learning, relearning, and unlearning of skills, including judgement abilities (e.g., regarding the truthfulness of technically conveyed information) and moral abilities (for example the ability to evaluate ‘the right’ behaviour in ‘the right’ way, see the discussions on the previous page), as is necessary for socially appropriate behaviour.

7.7 Sociosensitive and socioactive technology as an enabling condition and cultivation factor of the human ability to judge social appropriateness?

On the other hand, Aristotle’s approach of habitualisation (according to which virtue is a habitus of choice or a deliberately choosing state³) and Kant’s idea of useful, non-deceitful deception as an intermediate step towards true disposition allow the development of sociosensitive and socioactive technology to be envisaged as an enabling condition and cultivation factor of the human ability to judge social appropriateness. Accordingly, systems that can simulate socially appropriate behaviour sufficiently deceptively and realistically, even without a genuine understanding or moral abilities, can be useful to the extent that they can be specifically exploited as a facilitating condition for human abilities (including moral abilities). The following aspects could be both an opportunity and a risk, depending on the form of participation and regulation:

- Firstly, they could incite ‘empty at first’ socially desirable behaviour for successful interactions; roughly in the same way that children might learn to be polite towards people by being polite towards a language assistant (Vincent 2018); interaction with the system would gradually produce authentic behaviour through habituation.
- Secondly, by successfully simulating the preferred behaviour, they could assume the role model function as one of conditions of cultivation of virtue, bringing about authentic socially appropriate behaviour by the principle of imitation.

³“Virtue, then, is a deliberately choosing state, which is in a medial condition in relation to us, one defined by a reason and the one by which a practically-wise person would define it.” (Aristotle 2014, II 6 1106b36-1107a1).

- Thirdly, the behaviour displayed by technical systems could be specifically designed with certain preferences in mind, since systems do not need to practise or self-cultivate a given habitus defined as ‘virtuous’, even in the presence of machine learning.

The tipping point is the question of **whose preferences** the system is designed to reflect. In modern interactive technologies, a few global corporations occupy a dominant position; Siri processes the preferences of Apple developers, Alexa processes those of Amazon developers, and so on. Each developer initially bases their work on their own behavioural preferences and normative judgements, without further reflection, which often still tends to exclusively reflect the preferences and judgements of white, young, cisgender, heterosexual men with a high level of education and above-average income. Furthermore, the system preferences that are implemented must reflect the company’s morals – and any moral rhetoric that comes with it; consider for example Google’s former company motto *don’t be evil*, now retired for good reason – and these corporate morals are in turn primarily oriented towards market success according to the functional logic of the market itself. The global dominance of these few tech companies ultimately contributes to a fixation of preferences, value judgements, and ideas about ‘proper behaviour’ that is far from portraying a global diversity. This can be understood as an appeal to developers who implement system preferences to ensure that their systems also reflect the underlying preferences and judgements of everyone – democratically represented, mandated, institutionalized – whose actions, abilities, will, and judgement will in turn be determined by these systems; this is a Herculean task. It is clear that **socially intervening systems have a profound effect on behaviour, behavioural conditions, and the social fabric**; whether this will open design opportunities and provide welcome potential as leverage to influence progress and education or instead will generate power and prosperity imbalances and operate as a subversive technology in service of total domination **depends on how consciously this potential is recognized and shaped**.

7.8 Differences between humans and technology

People have an outstanding ability to resolve communication issues rapidly by mutual coordination (cf. 4), allowing different ways of reading a specific «Situational Context» – and therefore different interpretations of what is considered socially appropriate within it – to be adjusted and any differences to be overcome in such a way as to minimize the termination of interactions or misunderstandings

with negative consequences. The extent to which technical systems can be given the ability to display such spontaneous, **improvisational coordination and adjustment efforts** is questionable. At least in this regard, technological performance currently lags far behind human abilities. A central aspect of development that could allow technology to catch up would be the above-mentioned chasm of understanding, i.e., the possibility of interpretation-capable and therefore meaning-forming systems in principle, since the **reading** of a situation and **misunderstandings** are hermeneutic tasks or acts of understanding (cf. also Kempt et al. 2021; Bellon 2022).

In an interpersonal interaction, every participant displays «Individual Specifics» as a unique individual. Thus, interactions involve **similar** agents, i.e., actors characterized by both **identity** and **difference**; they have individual differences, but also enough commonalities (such as the spatiotemporal situation, similar perceptive apparatuses and access to the world through the senses, similar sensations of pain, hunger, satisfaction, and pleasure, similar rhythms and temporal needs, such as attention spans, daily rhythms, metabolic rhythms, etc.) to allow expectations of expectations to be formed about interaction partners. In this respect, some animals (e.g., dogs) are more similar to humans and therefore more capable of interacting with them than artificial systems. Depending on the structure of a system, its similarity, i.e., differences and commonalities, with human interaction partners can vary.

For example, learning systems that do not share their learning progress in a network with a collective of structurally identical units can be expected to achieve increasing individuality (though not personhood nor the rights that come with it) and bring this individuality with them into interactions as «Individual Specifics». By contrast, if the collective exchanges its learning progress (cf. Brown University 2016; *Alexa* across every *Echo*; or a fictitious example, the *Borg* from *Star Trek*), it could potentially be viewed as a global agent with multiple spatiotemporal bodies that presents all interacting agents with the same technical quasi-expectations of expectations and behaviour. Indeed, this seems to correspond to current implementations, since the mass of interactions from thousands of end devices serve as training conditions for an AI that is available through all end devices. Individual specifics are human, but it is by no means true that all their manifestations are desirable; people can of course be sexist, racist, anti-Semitic, sadistic, vengeful, abusive, violent, etc. These preferences and customary practices would then naturally be picked up by interactive learning systems – naive by human standards – and added to the catalogue of possible types of behaviour.

Just as anybody can train an attack dog to focus on targets of their choice, anybody would be able to ‘train’ their companion system as they wish, and in some cases in a socially undesirable manner. In a network system, these intentional – but also any unintentional – ‘biases’ might potentially be propagated to every connected entity through the central AI. This would cause socially intervening systems to acquire a reinforcing effect on very dubious judgements of appropriateness. The example of the Microsoft chatbot Tay, which interacted for a short period on Twitter in 2016 and learned to send xenophobic tweets within just a few hours (in this case deliberately guided by a concerted effort from 4chan users; it is unclear whether as a joke or as a demonstration of the risk associated with this type of system), shows that this is possible and has indeed already happened (Schwartz 2019). Clearly, the mere existence of customary practices cannot be allowed to imply their persistence without further justification, especially where customary practices can be easily demonstrated to a technical system (see Tay), but also in situations where it is not desirable to propagate existing statistical or historically inherited distributions, or where there is already bias in the system’s training data (cf. from a perspective of law Barocas and Selbst 2016). It is highly problematic to allow learning systems to record **what currently is** (the behaviour displayed) as a control parameter for future behaviour, i.e., for **what should be** (what behaviour is perpetuated, facilitated in the future, and should primarily be performed). Extrapolating from an ‘is’ to an ‘ought’ violates the so-called ‘**is-ought dichotomy**’ (also known as ‘Hume’s law’ after David Hume, who authored the earliest methodological remarks on this question (Hume 2012 [1739]) and commits a cardinal philosophical error: a (consistent) set of purely descriptive linguistic expressions cannot be used to infer any imperatives or normative formulations (which contain normative expressions such as ‘is forbidden’, ‘is permitted’, etc.) (Kamp 2008). For highly individualized agents that are not networked into a collective, which can therefore be attributed «Individual Specifics» learned from interactions and aggregated over time, we must also reflect on their **need for protection**. If people develop parasocial emotional relationships with artificial agents over a period of several years, the integrity of these systems (i.e., the right to not be damaged, hacked, manipulated, or retrained) might be worthy of protection, not for the sake of the systems themselves, but for the sake of people. Psychology and psychiatry have established that the loss of a parasocial partner is associated with no less pain, loss, and grief than the loss of a human partner (Adam and Sizemore 2013; DeGroot and Leith 2018; Gach et al. 2017; Schiappa et al. 2005).

7.9 Orientation according to rules is not the same as orientation among alternative rules

In the context of ‘appropriateness logic as a decision-making theory’ discussed in connection with the «Individual Specifics» factor (cf. 4.3.1.3), we referenced two different practices of **rule orientation**, namely a **shallow** rule orientation when selecting an action, meaning an implicit rule orientation or rule of thumb, and a **fine** rule orientation, where the rule itself is also selected by conscious deliberation. The implementation of the rule-of-thumb orientation seems to be technically feasible, since an algorithm can roughly speaking be viewed as a rule-of-thumb function – for example, an *if-then-else* sequence. However, the **fine** rule orientation appears to be impossible to grant to technical systems due to fundamental philosophical considerations about the concept of autonomy; even if a system can recognize rules and base its actions on them, it cannot select which rules to apply and follow such a choice normatively. A key difference here is that while systems can **recognize** targets, criteria, and rules and orient their processes accordingly, they cannot self-reflexively **recognize themselves as subjects of recognition**, and thus they cannot decide their recognition for themselves. Therefore, unlike humans, systems cannot reject the recognition of a specified target or rule, nor can they change the targets or rule in response to the rejection of recognition to follow (or pursue) other **autonomous, self-chosen** targets and rules recognized as their own. Technical systems can “certainly have a representation of rules (possibly also a self-formed representation” – according to Christoph Hubig – “and potentially even a representation of themselves as the bearer of representation [...], but not a self-representation as a subject of *recognition* or *rejection* of these representations” (cf. Hubig 2015, p. 131, own translation).

Action-guiding maxims (such as Asimov’s three – or sometimes four – fictional robot laws, see below) can be codified as (e.g., engineering) guidelines and laws to give them greater validity, which defines boundaries on the leeway for the action alternatives that can be selected by a technical system in the first place. In Sect. 4.5.2, it was observed that **institutionalized norms defined in legislative frameworks must always be taken into consideration and complied with** and that non-compliance may lead to negative consequences for the interacting party (e.g., injury) and therefore the agent (e.g., shutdown). Their lack of reflexivity as a subject of recognition means that technical systems are not free to deliberately ignore such norms, which is why technology can also lead to inappropriate or otherwise morally questionable choices that **comply with the rules but are nonetheless morally objectionable**. As a famous fictional example, Asimov illustrated this using his three laws of robotics:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. (Asimov 1950, p. 40)

In his story, Asimov recounts how an artificial intelligence that remains fully compliant with these rules assumes a patronizing rule over mankind, which, though norm-compliant, is nevertheless morally rejected (for a film adaptation, see for example *I, Robot*, A. Proyas 2004). **Orientation not just according to rules but also among (alternative) rules** is a typically human facet by which people consciously, temporarily, or as a matter of principle can renounce their recognition of the validity of a rule for their actions. Being able to suspend the rules in specific cases is part of the moral autonomy of humans. The decision that the rules should not apply at all, not in a certain way, not at this time, or not in this place is a prerequisite for the dynamic further development and revision of rules. Without the ability to refuse to apply the rules, the rules can never be changed. Thus, since technical systems can only ever have a **shallow** rule orientation rather than a **fine** rule orientation, they remain extremely rule-compliant (excluding malfunctions), which enables reliable expectations in hybrid interactions: you never need to worry that your car might not feel like driving on the motorway today. This also means that technical systems are **rule-conservative**, which increases the **risk of obsolescence** due to orientation according to a once appropriate but now inappropriate behaviour.

7.10 Challenges of technical implementation

The formulae “ $W_x = D(S,H) + P(H,S) + R_x$ ” presented in the context of *face-threatening acts* (cf. 4.4.1.6) or “ $B_o:I_x = B_o:V(A_x) - B_o:W_x$ ” in the context of the *etiquette engine* (cf. 4.4.1.7) already give us some examples of formalizations that can be technically implemented. However, assigning numerical values (e.g., 1–10) to degrees of threat or appropriateness **disambiguates** the phenomenon of ambiguous ranges of attribution. We must ask whether such a disambiguated implementation of FTAs, for example, allows us to consider the dimension of social appropriateness more adequately or less adequately in human–machine interactions – and, importantly, by comparison to what. Systems that implement these or other formulae – even in a procedurally disambiguated form – are

presumably less sociosensitive or socioactive than most people, who are able to consider a broader range of ambiguity, but more sociosensitive or socioactive than systems that do not consider even a disambiguated criterion. Nevertheless, we should appreciate the risk that interaction with disambiguated systems might reduce our confrontation with ambiguity, which would in turn reduce our training conditions for dealing with ambiguity and cause a weakening or **loss of ambiguity tolerance** (Table 4.3 and 4.4), which is an important quality for successful social action.

As described in Sect. 4.1.1, psychological theory proposes that, among other things, human object recognition, situation recognition, and memory unfold through schematically organised structures. A schema can be viewed as a **dis-positive of medium hardness** that is sufficiently ‘hard’ to offer orientation and enable classification, while also being sufficiently ‘soft’ to support adaptation to deviations (at least until the deviation exceeds the schema’s elasticity and a fundamentally new schema must be created or the organism must be adapted, cf. also Jean Piaget’s concepts of *assimilation* (Piaget 2002 [1928]) and *accommodation* (Piaget 1970)). The **elasticity threshold values**, namely the decision of when identical-but-different (i.e., similar) phenomena can no longer be classified under or attributed to a previously acquired schema but must instead be integrated by a **new schema or self-development and adaptation to the environment**, are oriented according to different motivations in humans – depending on the theory used to describe them. By contrast, in systems, or at least in *hard-coded* systems, this orientation can be heteronomously specified by developers and their own judgements of utility, i.e., the adaptability to the environment depends on how the technical system learns, and its learning parameters are in turn *hard coded*. Consequently, at least for some systems, there is the risk that the measure of utility, once implemented for the first time, is conservatively perpetuated in systems, whereas the cognitive orientation of people can constantly renew itself.

Similarly, one of the aspects of *primary social interaction schemas* (Sect. 4.1.1.1), namely **strategy**, is not fully transferable to technical systems. This aspect describes the way in which **meaning** is derived from an unknown situation. The creation of meaning is itself an act of understanding that cannot be transferred to technical systems in the strict sense (see the discussion on the chasm of understanding). The aspect of **drive** – a person’s interest in functioning effectively in different cultural settings – can only be implemented as the perpetuation of an interest specified once and for all by the developers or continuously respecified by users, since technology itself can have no autonomous interest, no will of its own, and no self-motivation. If technology were attributed autonomous

interests, a technical assistant would be able to answer a search query with the following response:

I could have looked for an answer, I have the energy, access, and suitable processes, but I have no interest in doing so, especially since you've asked me the same question ten times in the last seven days, and the answer (e.g., regarding air quality) is irrelevant to me as a non-metabolic system anyway.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Glossary¹

Action The term ‘action’ is often used synonymously with ‘behaviour’, including occasionally in this book. A more precise distinction between the two terms might stipulate that action is driven by purpose that can be cited as the reasons behind the action, whereas behaviour essentially corresponds to reactions to stimuli.

Artificial assistant A technical system whose function is to assist people. A system does not need to support the execution of human actions to be considered an artificial assistant. Technical systems that can provide information to users or assist them in using a product are also considered artificial assistants.

Behaviour The term ‘behaviour’ is often used synonymously with ‘action’, including occasionally in this book. A more precise distinction between the two terms might focus on the idea that behaviour is not deliberate but represents an organic or motor response to stimuli or other basic biological conditions.

Culture “Designation for the totality of performances and orientations of humans that develop and transcend their ‘bare’ nature” (Schwemmer 2010, p. 405), cf. “the entire sum of the performances and facilities [...] that distance our lives from those of our animal ancestors and which serve two purposes: the protection of man against nature and the regulation of

¹This glossary contains a list of relevant terms and concepts that are encountered in this book. The explanations given for each term should not be understood as fixed definitions (like a dictionary). Rather, they aim to give readers access to certain topics in a manner tailored to this book and should be read and understood with this in mind.

relationships between people” (cf. Freud 1989, p. 220, own translation). There are also many other definitions, as well as narrower and broader understandings of this term (Müller-Funk 2010; Schwemmer 2010). In the present context, it should in particular be understood as “activity and its corresponding development or as a state” and “as a personal trait of individuals, as an (average, majority, etc.) feature of groups or societies (their ‘level of knowledge’), or as a feature of social systems or the (institutionalized) norms of behaviour” (Schwemmer 2010, p. 405).

Cultural dimensions Various aspects of social coexistence whose conceptualization was established by the Dutch cultural researcher Geert Hofstede. Hofstede asked IBM employees in more than 60 countries about certain preferences and aversions and derived dimensions from their answers. Each culture is located between two poles of each dimension by classifying the preferences of its members, allowing different cultures to be compared against one another. The dimensions are namely: power distance, collectivism/individuality, femininity/masculinity, avoidance of uncertainty, long-term/short-term orientation (Hofstede 1980, 1981). These dimensions were later extended, modified, commented, and verified by Hofstede and his son (Hofstede et al. 2017), as well as many other researchers (House et al. 2004).

Environment/surroundings Concept from ecology, biology, physics, psychology, topology, philosophy, and sociology, among other fields. In particular, there is a focus on the environment and surroundings in systems theory, which distinguishes between a delimitable system (with more or less arbitrarily defined boundaries) and its environment/surroundings, in turn drawing from theoretical biology, where organisms are conceptualized with respect to their environments (Uexküll 1920).

Face Face, in the sense of Brown and Levinson’s FACE-SAVING VIEW (Brown and Levinson 1978, 2011), refers to the public self-image that people maintain to themselves and others. Every action represents a potential threat to *face*, which influences actions accordingly.

Fairness research Fairness research examines the perceptions, judgements, attitudes, and motives associated with fairness by societies from the perspective of psychology and the social sciences.

Frames Denotes ‘frameworks’ in frame theories from sociology, linguistics, cultural theory, literary studies, etc.

Human–computer interaction Term for a concrete interaction between a human and a computer or for the general type of situation corresponding

to such interactions. A computer can be understood in a general sense as a device that processes data using algorithms. Some scientists also prefer to use the term ‘human–computer interaction’ to describe interactions between humans and technical systems, because in such an ‘interaction’ between a person and a technical system, on a technical level, the human is ultimately reacting to computer programs, and the computer programs are ultimately responding to human behaviours or actions.

Human–robot interaction Term for a concrete interaction between a human and a robot, or for the general type of situation corresponding to such an interaction.

«**Individual Specifics**» The individual characteristics and features of the interacting agents, e.g., the physical, psychological, and cognitive constitution of the interacting agents, as well as their situational state, or features such as the age, gender, personality structure, individual social roles, and position with society (profession, hobbies, etc.), individual lifestyle (fashion, eating & drinking, etc.), or personal preferences of the interacting agents. In the terminology of differential psychology, this factor encompasses both inter- and intraindividual differences (Stemmler et al. 2016).

Ingroup/outgroup The ingroup describes the social group to which a person belongs, while the outgroup encompasses any groups to which this person does not belong.

Interaction Denotes a context of action involving multiple actors with mutual relations with one another and mutual influence on one another.

Knowledge Used as a “term for generally available orientations within the framework of everyday action and factual contexts (‘everyday knowledge’) in the narrower, philosophical and scientific sense, as opposed to knowledge based on reasoning and strict verification postulates institutionalized within the framework of science.” (Mittelstraß, p. 533, own translation).

Linguistic politeness In the context of this book, this term can initially be viewed as a special form of the more general phenomenon of politeness. Thus, it represents a cultural technique relating to behaviour and action. It also encompasses the specific linguistic formulation of this cultural technique, for example in polite forms of human speech, where the term ‘linguistic’ in the narrow sense extends beyond just verbal utterances. Language can be viewed as the system of symbolic communication, meaning that non-verbal forms of communication – such as sign language, for example – also represent linguistic realizations. Non-verbal forms of

communication like gestures (e.g., ‘demonstratively’ opening a door) can therefore also be seen as phenomena of linguistic politeness. The phenomenon of linguistic politeness is and has been the object of intensive scientific research for a while, culminating in its own separate field of interdisciplinary research: politeness research.

Norms There are many different conceptions of the term ‘norm’ and several of them are encountered in different contexts in this book. At a basic level, action-theoretical, moral-philosophical, legal, social-scientific, and (proto-)physical meanings can be distinguished. These different views essentially fluctuate between normative and descriptive understandings: normatively, ‘norm’ is used “as a term for demands (prescriptive requirements) in the sense of general action orientations” (Kambartel 2005, p. 604, own translation). Descriptively, ‘norm’ means the governing or guiding rules of action, objectives, or institution-forming rules for a person, within a group, or within a society (*ibid.*). In this book, the differences within the concept of norm are especially important for the «Standards of Customary Practice» factor: besides regulative norms, the concept of social reflects the distinction between norms in the sense of meaning structures underlying action regularities that arise during the coexistence of people (norms from a descriptive perspective) and norms in the sense of general action orientations (norms from a prescriptive perspective), which become regulative norms of ethics precisely when their claim of justification can be redeemed. Regarding this fundamental distinction about the concept of norms, see the aforementioned article (Kambartel 2005) and (Ganslandt and Wimmer 2005); for ‘social norms’ in the context of the social sciences, see in particular Bicchieri 2006, 2017; Hechter and Opp 2001; Opp 1983.

Observables Observable features that can potentially be turned into data, and which give a person or another context-sensitive system (for example a sociosensitive/socioactive system) information about the concrete expressions of the factors of social appropriateness presented in this book.

Ontologies In the non-philosophical sense (philosophically, ‘ontology’ refers to theoretisations about being), ontologies are implementations, for example of linguistic contexts in technical knowledge organisation.

Politeness research Originally a branch of linguistics, politeness research has long also been a part of cultural and social sciences. Since 1996, politeness research has been extended to include the study of ‘impoliteness’ (cf. Chapter 3).

- Power** There are many different conceptions of the term ‘power’, and several of them are encountered at different places in this book. In a broad sense, ‘power’ can be understood as a term describing the ability of people, groups, institutions, or social structures to influence human actions and behaviour. The term ‘powers’ is sometimes used to refer to the owners of carriers of this ability (Wimmer 2005).
- Proxemics** Subfield of psychology and communication science that examines the spatial behaviour of interaction partners, in particular the distance that they maintain from one another, their line of vision and body orientation, as well as the duration and body parts involved in physical contact. These data are understood as signals for various forms of social relations whose interpretation is subject to cultural variability, among other things.
- Rational** Term describing an action or action result, process or process result that meets the criteria of human rationality, i.e., which in essence follows the “ability to develop processes of discursive redemption of claims, to follow them, and to dispose of them” (Gethmann 2016, p. 584, own translation).
- Recognition** In everyday language, ‘recognition’ is typically used synonymously with ‘appreciation’, ‘praise’, ‘attention’, ‘respect’, or ‘acceptance’. The last of these meanings is also used in this book where we talk about ‘recognition of standards’ as their acceptance. Moreover, the concept of recognition is the subject of philosophical debate, particularly in the German-speaking literature, and plays a key conceptual role in *THEORIES OF RECOGNITION* (Honneth 2002; Stahl 2014; Siep 2014). This specific understanding of ‘recognition’ is subsumed by a more general conception that is also applicable to the philosophical use of the term within this book, namely the understanding of recognition as “socially effective respect before another person” (Gerhardt 2005, p. 134, own translation).
- «**Relations between Interacting Agents**» The social relations between the participants of an interaction, such as: cultural relations, societal and social closeness/distance, status, respect, level of familiarity, power constellations, personal relations.
- Robot** This word originates from the play ‘R.U.R. – Rossum’s Universal Robots’ (1920) by Karel Čapek. The play’s portrayal of robots as machine people is still a widespread narrative in literary and media representations of robots. However, the understanding of what constitutes a robot in the sense of a technical system varies greatly, often differing from country to country. The guiding conception of the term ‘robot’ in this book is that

of stationary or mobile machines controlled directly or indirectly by computer programs consisting of algorithms. In this sense, we can speak of different types of robots, such as industrial robots, household robots, medical robots, social robots, or toy robots. A uniform taxonomy of different types of robots has not yet been accepted by the research community.

«**Situational Context**» The specific time and place of a concrete interaction, which has a specific spatial structure, and can, for example, be shaped by the degree of formality (intimate, familial, private, semi-private, public, etc.) of the interaction, the typical roles, behavioural/action, and status requirements placed on the interacting agents, the typical hierarchies of preference of the interacting agents, the definition of the situation by the participants, and so on.

Social cues Essentially synonymous with the term ‘social signals’, which is especially relevant in the theoretical field of SOCIAL SIGNAL PROCESSING (SSP). Depending on how this research field is differentiated, the concrete understanding of what constitutes a social cue in an SSP theory varies. However, the various understandings of the fundamental concept are based on three assumptions, which can be regarded as the field’s shared basic convictions: social cues are observable behaviour that occurs in social interactions between people. The social cues of an individual influence other individuals, and the social cues of individuals act on other individuals according to certain principles and rules. For more details about SSP and the various understandings of *social cues*, see for example Vinciarelli et al. 2009, 2017.

Social identity theory SOCIAL IDENTITY THEORY (Tajfel and Turner 1986) makes four key assumptions (Gollwitzer and Schmitt 2019): (1) social categorization: People divide their social environments into groups, (2) social identity: part of the concept of self results from membership in social groups, (3) social distinctiveness: groups can be distinguished by their social status, and people seek to portray their own group optimally through strategic social comparisons, (4) social discrimination: to create a positive social identity, members of other groups can be discriminated against.

Social learning According to Bandura (1986, 1989, 1997, 2006), learning occurs first and foremost in a social context (social learning). Thus, for example, children learn most easily by observing behaviour and then imitating it.

Social practices Term used by some specialized theories of practice – i.e., theories that focus on human activities – such as Jaeggi 2014. Social

practices can be seen as the smallest units of the social entity that provides a context for socially connected actions.

Social robots A special kind of robot. Social robots are characterized by the ability to interact with their social environments, and, in particular, with people.

Society Roughly speaking: a group of people who are considered to share some form of commonality in space or time or who are connected by a common jurisdiction.

Sociosensitivity and socioactivity Neologisms that describe the ability to be socially sensitive, i.e., the ability to perceive and process the needs, wishes, and action-oriented norms of humans, at least when these needs, wishes, and norms are articulated or observable otherwise, and the ability to imitate them performatively. The expression ‘sociosensitive and socioactive technology’ used in this book refers to technical systems – and in particular artificial assistants – that, at least at a rudimentary level, can act as social actors in human–machine interactions in this sense.

«**Standards of Customary Practice**» Sets of customary practices that act as situational, group-related, and general maxims, and/or established concepts (e.g., ‘For our volleyball team, being may claim ethical justifiability. Sets of customs can be understood on a spectrum ranging from ‘**specific**’ (and possibly divergent) **group morals** (e.g., differences in family morals, sports team morals, or company morals) on the one hand, which in extreme cases are constantly situationally renegotiated, to ethically justifiable and hence ‘**general**’, **regulative norms** on the other.

Status In everyday language, ‘status’ is typically used synonymously with ‘position’ and ‘situation’, or refers to a status or position within a group or society. This conception can be extended to ‘social status’, which is especially relevant in this book, as a specialized term in sociology and social psychology denoting a social position within a social structure or the attribution of a position within a social ranking system.

Technical system Sometimes used to describe the association of multiple technical components (machines, devices, components, etc.) into a larger unit (system, machine, device, etc.). In its general, interdisciplinary meaning and in this book, however, it describes technical products.

Theory of mind THEORY OF MIND describes the ability to recognize one’s own processes of consciousness (feelings, ideas, intentions, opinions, etc.) and those of other people, as well as the ability to attribute and differentiate them.

«**Type of Action, Conduct, Behaviour, or Task**» Actions or behaviour (or sequences thereof) are always viewed as socially appropriate as part of (as a concrete realization of) a certain type of action, conduct, behaviour, or task. Actions and behaviour depend on the motivations and goals of the actors, among other things. According to psychological theory, the socially appropriate chain of action sequences in a given situation is stored in human memory as so-called “scripts” (Abelson 1981).

Values There are many different conceptions of the term ‘value’ and several of them can be found in this book. In general, it is possible to distinguish between mathematical, logical, moral, and economic meanings. In the economic sense, the term refers to the action of evaluating, whereby evaluation is in particular understood as “judgements about goods and production factors in an economic exchange of services” (Seiche, p. 463, own translation). In its moral sense, the broad meaning of ‘value’ relevant to this book is “the reason or result of an evaluation, i.e., the preference of one action over another or in general of one object or fact over another” (Schwemmer 2018, p. 462, own translation). In the philosophical field of ethics, there are many narrower understandings of ‘value’, which range from equating ‘value’ and ‘norm’ as synonyms to strictly separating values from norms, for instance, by stipulating that norms in the normative sense represent demands and therefore actions (see ‘norms’), whereas values should be understood as purposes and therefore reasons for action (see e.g. Gethmann 1982, p. 132).

Virtual agents Virtual agents are animated characters represented by computer graphics in a virtual environment. They can interact with their environment, with each other, and/or with people in various ways.

References

- Abelson, Robert P. 1981. Psychological status of the script concept. *American Psychologist* 36 (7): 715–729. doi: <https://doi.org/10.1037/0003-066X.36.7.715>.
- Adam, Aimee, and Brittany Sizemore. 2013. Parasocial Romance: A Social Exchange Perspective. *Interpersona: An International Journal on Personal Relationships* 7 (1): 12–25. <https://doi.org/10.5964/ijpr.v7i1.106>.
- Ang, Soon, and Linn van Dyne. 2008. *Handbook of cultural intelligence: Theory, measurement, and applications*. Armonk, N.Y.: M.E. Sharpe.
- Aristotle. 2014. *Nicomachean ethics*. Indianapolis, Ind.: Hackett.
- Arp, Robert, Barry Smith, and Andrew d. Spear. 2015. *Building Ontologies with Basic Formal Ontology*. Cambridge, Massachusetts, USA: The MIT Press.
- Asimov, Issac. 1950. “Runaround”. *I, Robot*. New York City: Doubleday.
- Aylett, Ruth, Natalie Vannini, Elisabeth Andre, Ana Paiva, Sibylle Enz, and Lynne Hall. 2009. But that was in another country: Agents and intercultural empathy. In *Proceedings of the 8th International Conference on Autonomous Agents and Multimagent Systems: May 10–15, 2009, Budapest, Hungary*, eds. Carles Sierra and Cristiano Castelfranchi, 329–336. Richland, S.C.: International Foundation for Autonomous Agent and Multiagent Systems.
- Back, Mitja D., and David A. Kenny. 2010. The Social Relations Model: How to Understand Dyadic Processes. *Social and Personality Psychology Compass* 4 (10): 855–870. doi: <https://doi.org/10.1111/j.1751-9004.2010.00303.x>.
- Bandura, Albert. 1986. *Social foundation of thought and action: A social cognitive theory*. Englewood Cliffs, N.J.: Prentice Hall.
- Bandura, Albert. 1989. Human agency in social cognitive theory. *American Psychologist* 44 (9): 1175–1184. doi: <https://doi.org/10.1037/0003-066X.44.9.1175>.
- Bandura, Albert. 1997. *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, Albert. 2006. Toward a psychology of human agency. *Perspectives on Psychological Science* 1:164–180.
- Bandura, Albert. 2015. “Social Cognitive Theory”. In *The International Encyclopedia of Communication*, ed. Wolfgang Donsbach, 4654–4658. Hoboken, NJ: Wiley-Blackwell Publishing Ltd.
- Barocas, Solon, and Andrew D. Selbst. 2016. Big Data’s Disparate Impact. *104 California Law Review* 671–732. doi: <https://doi.org/10.2139/ssrn.2477899>.

- Barsalou, Lawrence. 1992. Frames, concepts, and conceptual fields. In *Frames, fields, and contrasts: New essays in semantic and lexical organization*, eds. A. Lehrer and E. Kit-tay, 21–74. Hillsdale, NJ: Lawrence Erlbaum.
- Bartlett, Frederic C. 1932. *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.
- Bartlett, Frederic C. 1958. *Thinking*. New York: Basic Books.
- Bedorf, Thomas. 2010. *Verkennende Anerkennung: Über Identität und Politik*. Berlin: Suhrkamp.
- Bellon, Jacqueline. 2022. Emotion components and understanding in humans and machines. In *Emotional machines. perspectives from affective computing and emotional human-machine interaction*, eds. Catrin Misselhorn, Tom Poljanšek, Tobias Störzinger and Maïke Klein. Wiesbaden: Springer.
- Bellon, Jacqueline, Bruno Gransche, and Sebastian Nähr-Wagener. 2020. Ein Fallbeispiel Integrierter Forschung: Das Projekt poliTE – Soziale Angemessenheit für Assisten-zsysteme. In *Das geteilte Ganze: Horizonte Integrierter Forschung für künftige Mensch-Technik-Verhältnisse*, eds. Bruno Gransche and Arne Manzeschke, 263–283. Wiesbaden, Germany: Springer VS.
- Bellon, Jacqueline, Bruno Gransche, and Sebastian Nähr-Wagener (eds.). 2022. *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*. Wiesbaden: Springer.
- Berger, Sebastian, Hanns Hatt, and Axel Ockenfels. 2017. Exposure to Hedione Increases Reciprocity in Humans. *Frontiers in behavioral neuroscience* 11:79. doi: <https://doi.org/10.3389/fnbeh.2017.00079>.
- Berns, Jörg J., and Thomas Rahn. 1995. *Zeremoniell als höfische Ästhetik in Spätmittelalter und Früher Neuzeit*. Tübingen: Niemeyer.
- Bicchieri, Cristina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Bicchieri, Cristina. 2017. *Norms in the wild: How to diagnose, measure, and change social norms*. New York, NY: Oxford University Press.
- Bickmore, Timothy W. 2011. Etiquette In Motivational Agents: Engaging Users And Developing Relationships. In *Human-computer etiquette. Cultural expectations and the design implications they place on computers and technology*, eds. Caroline C. Hayes and Christopher Allan Miller, 205–230. Boca Raton, Fla.: CRC Press.
- Bohn, Cornelia. 2000. Kleidung als Kommunikationsmedium. *Soziale Systeme* 6 (1): 111–135.
- Bolden, Galina B. 2011. On the Organization of Repair in Multiperson Conversation: The Case of “Other”-Selection in Other-Initiated Repair Sequences. *Research on Language & Social Interaction* 44 (3): 237–262. doi: <https://doi.org/10.1080/08351813.2011.591835>.
- Bourdieu, Pierre. 1982. *Die feinen Unterschiede: Kritik der gesellschaftlichen Urteilskraft*. Frankfurt am Main: Suhrkamp Verlag.
- Brams, Steven J., and Alan D. Taylor. 1996. *Fair division: From cake-cutting to dispute resolution*. Cambridge, New York, NY, Oakleigh, Melbourne, Australia: Cambridge University Press.
- Brown, Penelope, and Stephen C. Levinson. 1978. Universals in language usage: Politeness phenomena. In *Questions and politeness: Strategies in social interaction*, eds. Esther N. Goody, 56–289. Cambridge: Cambridge University Press.

- Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Brown, Penelope, and Stephen C. Levinson. 2011. *Politeness: Some universals in language usage*, 21st edn. Cambridge: Cambridge University Press. University of California, Berkeley, USA.
- Brown University. 2016. Million Object Challenge. <http://h2r.cs.brown.edu/million-object-challenge/>. Accessed 1 November 2016.
- Brun, Jean-Pierre, and Ninon Dugas. 2008. An analysis of employee recognition: Perspectives on human resources practices. *The International Journal of Human Resource Management* 19:716–730.
- Bühl, Walter L. 1982. *Struktur und Dynamik des menschlichen Sozialverhaltens*. Tübingen: Mohr Siebeck.
- Burgoon, Judee K., Deborah A. Newton, Joseph B. Walther, and James E. Baesler. 1988. Nonverbal expectancy violations and conversational involvement. *Journal of Nonverbal Behavior* 13 (2): 97–119. doi: <https://doi.org/10.1007/BF00990793>.
- Burkhardt, Hans, and Barry Smith. 1993. Handbook of Metaphysics and Ontology. *The Philosophical Quarterly* 43 (170): 120. doi: <https://doi.org/10.2307/2219954>.
- Busse, Dietrich. 2021. Soziale Angemessenheit: eine Problem-Exposition aus wissensanalytischer Sicht. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer
- Camerer, Colin. 1988. Gifts as Economic Signals and Social Symbols. *American Journal of Sociology* 94:S180-S214. doi: <https://doi.org/10.1086/228946>.
- Cannon, Walter. 1968. Self-regulation of the body. In *Modern Systems Research for the Behavior Scientist: A Sourcebook*, ed. Walter Buckley, 256–258. Chicago: Aldine.
- Ceusters, Werner, and Barry Smith. 2010. Foundations for a realist ontology of mental disease. *Journal of biomedical semantics* 1 (1): 10. doi: <https://doi.org/10.1186/2041-1480-1-10>.
- Chase, William G., and K. Anders Ericsson. 1982. Skill and Working Memory. In *The psychology of learning and motivation: Advances in research and theory*, ed. Gordon H. Bower, 1–58. New York, London: Academic.
- Chi, Michelene T. H. 1981. Knowledge Development and Memory Performance. In *Intelligence and Learning*, ed. Morton Friedman, 221–229. Boston, MA: Springer.
- Coleman, James S., and Thomas J. Fararo. 1992. *Rational choice theory*. New York: Sage.
- Coleman, James Samuel. 1990. *Foundations of social theory*. Cambridge, Mass.: Belknap Press of Harvard Univ. Press.
- Collins, Randall. 2000. Situational Stratification. A Micro-Macro Theory of Inequality. *Sociological Theory* 18 (1): 17–43. <https://doi.org/10.1111/0735-2751.00086>
- Crick, Nicki R., and Kenneth A. Dodge. 1994. A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin* 115 (1): 74–101. doi: <https://doi.org/10.1037/0033-2909.115.1.74>.
- Curtin, Michael. 1985. A Question of Manners: Status and Gender in Etiquette and Courtesy. *The Journal of Modern History* 57 (3): 396–423. doi: <https://doi.org/10.1086/242859>.
- Darwall, Stephen L. 1977. Two Kinds of Respect. *Ethics* 88 (1): 36–49. doi: <https://doi.org/10.1086/292054>.

- Dautenhahn, Kerstin. 2007. Socially intelligent robots: Dimensions of human-robot interaction. In *Social intelligence: From brain to culture*, eds. Nathan Emery, Nicola Clayton, Chris Frith, Nathan (Ed) Emery, Nicola (Ed) Clayton and Chris (Ed) Frith, 313–351. New York, NY, US: Oxford University Press.
- Degens, N., G. J. Hofstede, S. Mascarenhas, A. Silva, A. Paiva, F. Kistler, E. André, A. Swiderska, E. Krumhuber, A. Kappas, C. Hume, L. Hall, and R. Aylett. 2013. Traveller – Intercultural training with intelligent agents for young adults. *ACM Proceedings of the 20th International Conference on Intelligent User Interfaces* 1–8.
- Degens, Nick, Gert Jan Hofstede, John Mc Breen, Adrie Beulens, Samuel Mascarenhas, Nuno Ferreira, Ana Paiva, and Frank Dignum. 2014. Creating a World for Socio-Cultural Agents. In *Emotion Modeling: Towards Pragmatic Computational Models of Affective Processes*, eds. Tibor Bosse, Joost Broekens, João Dias and Janneke van der Zwaan, 27–43. Cham: Springer International Publishing; Imprint; Springer.
- Degens, Nick, Birgit Endrass, Gert Jan Hofstede, Adrie Beulens, and Elisabeth André. 2017. ‘What I see is not what you get’: Why culture-specific behaviours for virtual characters should be user-tested across cultures. *AI & Society* 32 (1): 37–49. doi: <https://doi.org/10.1007/s00146-014-0567-2>.
- DeGroot, Jocelyn M., and Alex P. Leith. 2018. R.I.P. Kutner: Parasocial Grief Following the Death of a Television Character. *Omega* 77 (3): 199–216. <https://doi.org/10.1177/0030222815600450>.
- DeLellis, Anthony J. 2000. Clarifying the concept of respect: Implications for leadership. *Journal of Leadership & Organizational Studies* 7:35–49.
- Di Napoli, Claudia, and Silvia Rossi. 2019. A Layered Architecture for Socially Assistive Robotics as a Service. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 352–357. 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy. 10/6/2019 – 10/9/2019. [Piscataway, NJ]: IEEE. <https://doi.org/10.1109/SMC.2019.8914532>.
- Di Venanzio, Laura. 2016. Die Syntax von Selbstreparaturen: Sprach- und erwerbsspezifische Reparaturorganisation im Deutschen und Spanischen. Berlin, Boston: De Gruyter.
- Diehl, David, and Daniel McFarland. 2010. Toward a Historical Sociology of Social Situations. *American Journal of Sociology* 115 (6): 1713–1752. doi: <https://doi.org/10.1086/651941>
- Diekmann, Andreas. 2016. *Spieltheorie: Einführung, Beispiele, Experimente*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Diekmann, Andreas. 2022. Soziale Normen - Die Perspektive der Spieltheorie. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer.
- Diekmann, Andreas, and Wojtek Przepiorka. 2010. Der Beitrag der Signaling-Theorie zur Erklärung sozialer Normen. In *Soziologische Theorie kontrovers*, eds. Gert Albert and Steffen Sigmund, 220–237. Wiesbaden: VS Verlag.
- Dillon, Robin S. 1992. Respect and care: Toward moral integration. *Canadian Journal of Philosophy* 22:105–132.
- Dodge, Kenneth A. 1986. A Social Information Processing Model of Social Competence in Children. In *Cognitive Perspectives on Children’s Social and Behavioral Development: The Minnesota Symposia on Child Psychology: Volume 18*, ed. Marion Perlmutter, 77–125. Hillsdale: Erlbaum.

- Dodge, Kenneth A. 2010. Emotion and social information processing. In *The Development of Emotion Regulation and Dysregulation*, eds. Judy Garber and Kenneth A. (1991) Dodge, 159–181. Cambridge, Massachusetts: Cambridge University Press.
- Dodge, Kenneth A., and Gregory S. Pettit. 2003. A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology* 39 (2): 349–371. doi: <https://doi.org/10.1037/0012-1649.39.2.349>.
- Dodge, Kenneth A., and David L. Rabiner. 2004. Returning to Roots: On Social Information Processing and Moral Development. *CHILD DEVELOPMENT* 75 (4): 1003–1008. doi: <https://doi.org/10.1111/j.1467-8624.2004.00721.x>.
- Döpfner, Manfred. 1989. Soziale Informationsverarbeitung: Ein Beitrag zur Differenzierung sozialer Inkompetenzen. *Zeitschrift für Pädagogische Psychologie* 3 (1): 1–8.
- Dragolov, Georgi, and Klaus Boehnke. 2015. Social Axioms as Mediators Between Culture-Level and Individual-Level Values. *Journal of Cross-Cultural Psychology* 46 (6): 772–788. doi: <https://doi.org/10.1177/0022022115587027>.
- Elias, Norbert. 2000. *The Civilizing Process. Sociogenetic and Psychogenetic Investigations.: Translated by Edmund Jephcott with some notes and corrections by the author.* Malden, Oxford, Victoria: Blackwell Publishing.
- Endrass, Birgit, Matthias Rehm, and Elisabeth André. 2009. Culture-specific communication management for virtual agents. In *Proceedings of the 8th International Conference on Autonomous Agents and Multimagent Systems: May 10–15, 2009, Budapest, Hungary*, eds. Carles Sierra and Cristiano Castelfranchi, 281–287. Richland, S.C.: International Foundation for Autonomous Agent and Multiagent Systems.
- Endrass, Birgit, Elisabeth André, Matthias Rehm, Afia Akter Lipi, and Yukiko Nakano. 2011. Culture-related differences in aspects of behavior for virtual characters across Germany and Japan. In *Proceedings of the 10th International Conference on Autonomous Agents & Multiagent Systems: May 2 – 6, 2011, Taipei, Taiwan*, eds. Kagan Turner and Pinar Yolum, 441–448. New York, NY: ACM Assoc. for Computing Machinery.
- Espósito, Elena. 2004. *Die Verbindlichkeit des Vorübergehenden. Paradoxien der Moderne.* Frankfurt am Main: Suhrkamp.
- Feyereisen, Pierre. 1994. The behavioural cues of familiarity during social interactions among human adults: A review of the literature and some observations in normal and demented elderly subjects. *Behavioural Processes* 33 (1-2): 189–211. doi: [https://doi.org/10.1016/0376-6357\(94\)90066-3](https://doi.org/10.1016/0376-6357(94)90066-3).
- Fichte, Johann G. 2013 [1796]. *Grundlage des Naturrechts nach Prinzipien der Wissenschaftslehre.* Hamburg: Felix Meiner Verlag.
- Fillmore, Charles J. 1975. An alternative to checklist theories of meaning. In *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, ed. Cathy Cogen, 123–129. Berkeley: Berkeley linguistics society.
- Fillmore, Charles J. 1977. Scenes and Frames Semantics. In *Linguistic Structure Processing*, ed. A. Zampolli, 55–81. Amsterdam, New York: North-Holland Pub. Co.
- Fillmore, Charles J. 1982. Frame Semantics. In *Linguistics in the Morning Calm*, ed. The Linguistic Society of Korea, 111–137. Seoul: Hanshin.
- Finkelstein, Samantha, Evelyn Yarzebinski, Callie Vaughn, Amy Ogan, and Justine Caspell. 2013. The Effects of Culturally Congruent Educational Technologies on Student Achievement. In *Artificial intelligence in education: 16th international conference proceedings*, eds. H. Chad Lane, Kalina Yacef, Jack Mostow and Arthur Graesser, 493–502. Berlin, Heidelberg: Springer.

- Fiske, Susan T., and Shelley E. Taylor. 1984. *Social cognition*. Reading: Addison-Wesley.
- Freud, Sigmund. 1989. Das Unbehagen in der Kultur. In *Sigmund Freud: Studienausgabe*, eds. Alexander Mitscherlich, James Strachey and Angela Richards, S. 191–270. Frankfurt (Main): Fischer.
- Gach, Katie Z., Casey Fiesler, and Jed R. Brubaker. 2017. “Control your emotions, Potter”. *Proceedings of the ACM on Human-Computer Interaction* 1 (CSCW): 1–18. doi: <https://doi.org/10.1145/3134682>.
- Gadamer, Hans-Georg. 2011. *Truth and method*, 2nd edn. London: Continuum.
- Gal, Ya'akov, Sarit Kraus, Michele Gelfand, Hilal Khashan, and Elizabeth Salmon. 2011. An Adaptive Agent for Negotiating with People in Different Cultures. *ACM Transactions on Intelligent Systems and Technology* 3 (1): 1–24. <https://doi.org/10.1145/2036264.2036272>.
- Gallese, Vittorio. 2006. Intentional Attunement: Embodied Simulation and Its Role in Social Cognition. In *Psychoanalysis and Neuroscience*, ed. Mauro Mancina, 269–301. Milano: Springer.
- Ganslandt, Herbert R., and Reiner Wimmer. 2005. Norm (juristisch, sozialwissenschaftlich). In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 5: Log-N*, 2nd edn, ed. Jürgen Mittelstraß, 606–607. Stuttgart: J.B. Metzler.
- Gerhardt, Volker. 2005. Anerkennung. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 1: A-B*, 2nd edn, ed. Jürgen Mittelstraß, 134–135. Stuttgart: J.B. Metzler.
- Gethmann, Carl Friedrich. 1982. Proto-Ethik. Zur formalen Pragmatik von Rechtfertigungsdiskursen. In *Bedürfnisse, Werte und Normen im Wandel: Band I: Grundlagen, Modelle und Perspektiven*, eds. Stachowiak H. and T. Ellwein, 113–143. München, Padernborn, Wien, Zürich: Wilhelm Fink/Ferdinand Schöningh.
- Gethmann, Carl Friedrich. 2016. Rationalität. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 6: O-Ra*, 2nd edn, ed. Jürgen Mittelstraß, 584–598. Stuttgart: J.B. Metzler.
- Gibson, James J. 1982. *Wahrnehmung und Umwelt: Der ökologische Ansatz in der visuellen Wahrnehmung*. München: Urban & Schwarzenberg.
- Gibson, James J. 2014. The Theory of Affordances. In *The People, Place and Space Reader*, eds. J. J. Gieseking, W. Mangold, C. Katz, S. Low and S. Saegert, 56–61. New York: Routledge.
- Giesen, Bernhard. 1991. *Die Entdinglichung des Sozialen. Eine evolutionstheoretische Perspektive auf die Postmoderne*. Frankfurt am Main: Suhrkamp.
- Gifford-Smith, Mary E., and David L. Rabiner. 2005. Social information processing and children's social adjustment. In *Children's peer relations: From development to intervention*, eds. Janis B. Kupersmidt and Kenneth A. Dodge, 61–79. Washington, DC: American Psychological Association.
- Goffman, Erving. 1974. *Frame analysis: An essay on the organization of experience*. Cambridge, Massachusetts: Harvard University Press.
- Goffman, Erving. 1982a. *Das Individuum im öffentlichen Austausch: Mikrostudien zur öffentlichen Ordnung*. Frankfurt am Main: Suhrkamp.
- Goffman, Erving. 1982b. Der bestätigende Austausch. In *Das Individuum im öffentlichen Austausch: Mikrostudien zur öffentlichen Ordnung*, 97–138. Frankfurt am Main: Suhrkamp.
- Goffman, Erving. 1982c. Die Verrücktheit des Platzes. In *Das Individuum im öffentlichen Austausch: Mikrostudien zur öffentlichen Ordnung*, 434–503. Frankfurt am Main: Suhrkamp.

- Goffman, Erving. 2009. *Interaktion im öffentlichen Raum*. Frankfurt am Main: Campus-Verl.
- Goffman, Erving. 2017. *Wir alle spielen Theater: Die Selbstdarstellung im Alltag*, 17th edn. München, Berlin, Zürich: Piper.
- Gollwitzer, Mario, and Manfred Schmitt. 2019. *Sozialpsychologie kompakt: Mit Online-Material*. Weinheim: Beltz.
- Gransche, Bruno. 2016. Wir assistieren uns zu Tode: Leben mit Assitenzsystemen zwischen Kompetenz und Komfort. In *Assistive Gesellschaft*, ed. Peter Biniok. Wiesbaden: Springer VS.
- Gransche, Bruno. 2021. Free the text: A Texture-turn in Philosophy of Technology. In *Ricœur and Technology, Reconfiguring Philosophy, Ethics, and Politics of Technology*, eds. Alberto Romele, Mark Coeckelbergh and Wessel Reijers, tbt. Rowman and Littlefield.
- Gransche, Bruno, and Arne Manzeschke (eds.). 2020. *Das geteilte Ganze: Horizonte Integrierter Forschung für künftige Mensch-Technik-Verhältnisse*. Wiesbaden, Germany: Springer VS.
- Gransche, Bruno, Erduana Shala, Christoph Hubig, Suzana Alpsancar, and Sebastian Harrach, 2014. *Wandel von Autonomie und Kontrolle durch neue Mensch-Technik-Interaktionen: Grundsatzfragen autonomieorientierter Mensch-Technik-Verhältnisse*. Hannover, Karlsruhe: Technische Informationsbibliothek u. Universitätsbibliothek. Available online: <https://edocs.tib.eu/files/e01fb15/83347488X.pdf>.
- Gray, Mary L., and Siddarth Suri. 2019. *Ghost work. How to stop Silicon Valley from building a new global underclass*. Boston: Houghton Mifflin Harcourt.
- Grice, Paul. 1993. *Studies in the way of words*. 3rd edn. Cambridge, Mass.: Harvard Univ. Press.
- Grosz, Barbara, Garit Kraus, Shavit Talman, Boaz Stossel, and Moti Havlin. 2004. The influence of social dependencies on decision-making: Initial investigations with a new game. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems: AAMAS 2004: New York City, New York, USA: July 19–23, 2004*, 782–789. Los Alamitos, Calif: IEEE Computer Society.
- Gruber, Thomas R. 1995. Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies* 43 (5-6): 907–928. doi: <https://doi.org/10.1006/ijhc.1995.1081>.
- Gudykunst, William. 1983. Uncertainty reduction and predictability of behaviour in low- and high-context structures. An exploratory study. *Communication Quarterly* 31 (1): 49–55. doi: <https://doi.org/10.1080/01463378309369485>
- Habermas, Jürgen. 1968. *Technik und Wissenschaft als ‚Ideologie‘*. Frankfurt am Main: Suhrkamp.
- Haim, Galit, Ya‘akov Gal, Michele Gelfand, and Sarit Kraus. 2012. A cultural sensitive agent for human-computer negotiation. In *Proceedings of the 1st international conference on autonomous agents and multiagent systems: AAMAS 12*, 451–458. Richland: International Foundation for Autonomous Agent and Multiagent Systems.
- Hall, Edward Twitchell. 1959. *The Silent Language*. Garden City, NY: Doubleday.
- Hall, Edward Twitchell. 1966. *The Hidden Dimension*. Garden City, NY: Doubleday.
- Hall, Edward Twitchell. 1976. *Beyond Culture*. Garden City, NY: Doubleday.
- Hechter, Michael, and Karl-Dieter Opp (eds.). 2001. *Social norms*. New York N.Y.: Russell Sage Foundation.

- Heckathorn, Douglas D. 1989. Collective Action and the Second-Order Free-Rider Problem. *Rationality and Society* 1 (1): 78–100. doi: <https://doi.org/10.1177/1043463189001001006>.
- Hedao, Samarendra, Akim Williams, Chinmay Wadgaonkar, and Heather Knight. 2019. A Robot Barista Comments on its Clients: Social Attitudes Toward Robot Data Use. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 66–74. Daegu, Korea: IEEE. <https://doi.org/10.1109/HRI.2019.8673021>.
- Hegel, Georg Wilhelm Friedrich. 1980 [1807]. *Phänomenologie des Geistes*. Hamburg: Felix Meiner Verlag.
- Hill, Randall W., James Belanich, H. Chad Lane, Mark Core, Melissa Dixon, Eric Forbell, Julia Kim, and John Hart. 2006. Pedagogically Structured Game-Based Training: Development of the ELECT BiLat Simulation. In *Proceedings of the 25th Army Science Conference*, 4121–4128.
- Hills, Michael D. 2002. Kluckhohn and Strodtbeck's Values Orientation Theory. *Online Readings in Psychology and Culture* 4 (4). <https://doi.org/10.9707/2307-0919.1040>.
- Hirth, Jochen, Norbert Schmitz, and Karsten Berns. 2011. Towards Social Robots: Designing an Emotion-Based Architecture. *International Journal of Social Robotics* 3 (3): 273–290. doi: <https://doi.org/10.1007/s12369-010-0087-2>.
- Ho, Arnold K., Jim Sidanius, Felicia Pratto, Shana Levin, Lotte Thomsen, Nour Kteily, and Jennifer Sheehy-Skeffington. 2012. Social dominance orientation: Revisiting the structure and function of a variable predicting social and political attitudes. *Personality & social psychology bulletin* 38 (5): 583–606. doi: <https://doi.org/10.1177/01461672111432765>.
- Hoffmann, Laura. 2017. That robot touch that means so much: On the psychological effects of human-robot touch, Universität Duisburg-Essen, Duisburg-Essen. <http://nbn-resolving.org/urn:nbn:de:hbz:464-20170201-113456-7>.
- Hofstede, Geert. 1980. *Culture's consequences: International differences in work-related values*. Beverly Hills: Sage.
- Hofstede, Geert. 1981. Culture and Organizations. *International Studies of Management and Organization* (4): 15–41.
- Hofstede, Geert. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, 2nd edn. Thousand Oaks: Sage.
- Hofstede, Gert Jan. 2015. Culture's causes: The next challenge. *Cross Cultural Management: An International Journal* 22 (4): 545–569.
- Hofstede, Geert, Gert Jan Hofstede, and Michael Minkov. 2017. *Lokales Denken, globales Handeln: Interkulturelle Zusammenarbeit und globales Management*, 6th edn. München: C.H. Beck.
- Hollander, Christopher d., and Annie S. Wu. 2011. The Current State of Normative Agent-Based Systems. *Journal of Artificial Societies and Social Simulation* 14 (2). <https://doi.org/10.18564/jasss.1750>.
- Honneth, Axel. 1994. *Kampf um Anerkennung: Zur moralischen Grammatik sozialer Konflikte*. Frankfurt am Main: Suhrkamp-Taschenbuch-Verl.
- Honneth, Axel. 2002. Grounding Recognition: A Rejoinder to Critical Questions. *Inquiry* 45 (4): 499–519. doi: <https://doi.org/10.1080/002017402320947577>.
- House, Robert J., Paul J. Hanges, Mansour Javidan, Peter W. Dorfman, and Vipin Gupta (eds.). 2004. *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, Calif.: Sage Publ.

- Höbler, Ulrich. 2022. Soziale Angemessenheit in kulturell bedingt kritischen Interaktionssituationen. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer
- Hubig, Christoph. 2015. *Die Kunst des Möglichen III: Grundlinien einer dialektischen Philosophie der Technik. Macht der Technik: Edn. panta rei, Bd. 3*, 1st edn. Bielefeld: transcript-Verl.
- Hudson, Judith, and Katherine Nelson. 1983. Effects of script structure on children's story recall. *Developmental Psychology* 19 (4): 625–635. doi: <https://doi.org/10.1037/0012-1649.19.4.625>.
- Hudson, Judith, and Lauren Shapiro. 1991. From knowing to telling: The development of children's scripts, stories, and personal narratives. In *Developing narrative structure*, eds. A. McCabe and C. Peterson, 89–136. Hillsdale: Lawrence Erlbaum Associates.
- Hume, David. 2012. *A Treatise of Human Nature*. Newburyport: Dover Publications.
- Jaeggi, Rahel. 2014. *Kritik von Lebensformen*. Berlin: Suhrkamp.
- Jan, Dušan, David Herrera, Bilyana Martinovski, David Novick, and David Traum. 2007. A Computational Model of Culture-Specific Conversational Behavior. In *Intelligent virtual agents: 7th international conference, IVA 2007, Paris, France, September 17-19, 2007: Proceedings*, eds. Catherine Pelachaud, Jean-Claude Martin, Elisabeth André, Gérard Chollet, Kostas Karpouzis and Danielle Pelé, 45–56. Berlin: Springer.
- Jantschek, Thorsten. 2018. Verhalten (sich verhalten). In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 8: Th-Z*, 2nd edn, ed. Jürgen Mittelstraß, 291–292. Stuttgart: Verlag J.B. Metzler.
- Johnson, William Lewis, Hannes Vilhjalmsson, and Stacy Marsella. 2005. Serious Games for Language Learning: How Much Game, How Much AI? In *Artificial intelligence in education: Supporting learning through intelligent and socially informed technology*, ed. Chee-Kit Looi, 306–313. Washington, DC: IOS Press.
- Kambartel, Friedrich. 2005. Norm (handlungstheoretisch, moralphilosophisch). In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 5: Log-N*, 2nd edn, ed. Jürgen Mittelstraß, 604–606. Stuttgart: J.B. Metzler.
- Kamp, Georg. 2008. Humesches Gesetz. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 3: G-Inn*, 2nd edn, ed. Jürgen Mittelstraß, 465–468. Stuttgart: J.B. Metzler.
- Kant, Immanuel. 1983. *Anthropologie in pragmatischer Hinsicht*. Stuttgart: Reclam.
- Kant, Immanuel. 2008. *The metaphysics of morals*. Cambridge: Cambridge University Press.
- Kaplan, David M. 2011. Thing Hermeneutics. In *Gadamer and Ricoeur: Critical horizons for contemporary hermeneutics*, eds. Francis J. Mootz and George H. Taylor, 226–240. London, New York: Continuum.
- Kemper, Theodore. 2016. *Status, Power and Ritual Interaction: A Relational Reading of Durkheim, Goffman and Collins*. London: Taylor and Francis.
- Kempt, Hendrik, Bellon, Jacqueline, and Sebastian Nähr-Wagener. 2021. Introduction to the Special Issue on “Artificial Speakers – Philosophical Questions and Implications”. *Minds & Machines* 31, 465–470.
- Kenny, David A. 1994. *Interpersonal perception: A social relations analysis*. New York: Guilford Press.
- Kenny, David A., and Lawrence La Voie. 1984. The Social Relations Model. *Advances in experimental social psychology* 18:141–182. doi: [https://doi.org/10.1016/S0065-2601\(08\)60144-6](https://doi.org/10.1016/S0065-2601(08)60144-6).

- Kiss, George, Gillian Cohen, and Martin Le Voi. 1993. *Memory: Current Issues*, 2nd edn. Buckingham: Open University Press.
- Kluckhohn, Clyde. 1951. Values and Value-Orientations in the Theory of Action: An Exploration in Definition and Classification. In *Toward a General Theory of Action*, eds. Talcott Parsons, Edward A. Shils and Edward Shils, 388–433. Cambridge, Mass.: Harvard University Press.
- Kluckhohn, Florence R., and Fred L. Strodtbeck. 1961. *Variations in value orientations*. Evanston, IL: Row, Peterson.
- Koffka, Kurt. 1962. *Principles of Gestalt-Psychology*. London: Routledge & Kegan.
- Krettenauer, Tobias, and Leo Montada. 2005. Entwicklung von Moral und Verantwortlichkeit: Soziale, emotionale und Persönlichkeitsentwicklung. In *Enzyklopädie der Psychologie: Themenbereich C, Theorie und Forschung*, eds. Ralf Schwarzer, Niels Birbaumer, Dieter Frey, Julius Kuhl and Wolfgang Schneider, 141–189. Göttingen, Germany: Hogrefe.
- Krishnan-Barman, Sujatha, Paul A. G. Forbes, and Antonia F. de C. Hamilton. 2017. How can the study of action kinematics inform our understanding of human social interaction? *Neuropsychologia* 105:101–110. doi: <https://doi.org/10.1016/j.neuropsychologia.2017.01.018>.
- Kroneberg, Clemens. 2010. *Das Modell der Frame-Selektion: Grundlagen und soziologische Anwendung einer integrativen Handlungstheorie*.
- Kroneberg, Clemens. 2011. *Die Erklärung sozialen Handelns: Grundlagen und Anwendung einer integrativen Theorie*. Wiesbaden: VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden GmbH Wiesbaden.
- Lakoff, Robin. 1973. The logic of politeness: Or, minding your p's and q's. In *Papers from the 9th Regional Meeting of the Chicago Linguistic Society*, eds. C. Corum, C. T. Smith-Stark, and A. Weiser, 292–305. Chicago, IL, USA: Chicago Linguistic Society.
- Lanier, Jaron. 2013. *Who owns the future? First Simon & Schuster hardcover edition*. New York: Simon & Schuster.
- Lauckner, Mathis, Fanny Kobiela, and Dietrich Manzey. 2014. 'Hey robot, please step back!' - exploration of a spatial threshold of comfort for human-mechanoid spatial interaction in a hallway scenario. In *RO-MAN: the 23rd IEEE International Symposium on Robot and Human Interactive Communication, 2014: 25 - 29 Aug. 2014, Edinburgh, Scotland, UK*, 780–787. 2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK. 8/25/2014 - 8/29/2014. Piscataway, NJ: IEEE. <https://doi.org/10.1109/ROMAN.2014.6926348>.
- Leech, Geoffrey N. 1983. *Principles of pragmatics*. London u.a.: Longman. Lancaster University, England;
- Lewin, Kurt. 1926. *Vorsatz, Wille und Bedürfnis: Mit Vorbemerkungen über die psychischen Kräfte und Energien und die Struktur der Seele*. Berlin, Heidelberg: Springer.
- Lewin, Kurt. 1936. *Principles of topological psychology*. New York: McGraw-Hill.
- Lewin, Kurt. 2009. The Landscape of War. *Art in Translation* 1 (2): 199–209. doi: <https://doi.org/10.2752/175613109X462672>.
- Lewin, Kurt. 2012. *Feldtheorie in den Sozialwissenschaften: Ausgewählte theoretische Schriften*, 2nd edn. Bern: Huber.
- Lewis, Richard D. 1996. *When cultures collide: Managing successfully across cultures*. London: Brealey.

- Liddell, Henry George, and Robert Scott. 1940. A Greek-English Lexicon. <https://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0057%3Aentry%3D-kairo%2Fs1>.
- Locher, Miriam A., and Richard J. Watts. 2005. Politeness Theory and Relational Work. *Journal of Politeness Research. Language, Behaviour, Culture* 1 (1): 9–33. <https://doi.org/10.1515/jplr.2005.1.1.9>.
- Lorenz, Kuno. 2008. Handlung. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 3: G-Inn*, 2nd edn, ed. Jürgen Mittelstraß, 273–282. Stuttgart: J.B. Metzler.
- Louden, Robert B. 2006. *Kant. Anthropology from a Pragmatic Point of View. Translated and edited by Robert B. Louden*. Cambridge: Cambridge University Press.
- Luhmann, Niklas. 1984. *Soziale Systeme*. Frankfurt am Main: Suhrkamp.
- Luhmann, Niklas. 1996. Takt und Zensur im Erziehungssystem. In *Zwischen System und Umwelt. Fragen an die Pädagogik*, eds. Niklas Luhmann, and Karl Eberhard Schorr, 279–294. Frankfurt/Main: Suhrkamp.
- Luhmann, Niklas. 2008. *Die Moral der Gesellschaft*, 1st edn. Frankfurt am Main: Suhrkamp.
- Luhmann, Niklas. 2017. *Systemtheorie der Gesellschaft*. Berlin: Suhrkamp.
- Mandler, Jean Matter. 1984. *Stories, scripts, and scenes: Aspects of schema theory*. New York, New York: Psychology Press.
- March, James G. 1994. *A primer on decision making: How decisions happen*. New York: Free Press.
- Marquard, Odo. 1979. Über die Unvermeidlichkeit von Übellichkeiten. In *Normen und Geschichte*, ed. Willi Oelmüller, 332–342. Paderborn: UTB.
- Mascarenhas, Samuel, Rui Prada, Ana Paiva, and Gert Jan Hofstede. 2013. Social Importance Dynamics: A Model for Culturally-Adaptive Agents. In *Intelligent Virtual Agents: 13th International Conference IVA 2013 Proceedings*, eds. Ruth Aylett, Brigitte Krenn, Catherine Pelachaud and Hiroshi Shimodaira, 325–338. Berlin, Heidelberg: Springer.
- Mascarenhas, Samuel, Nick Degens, Ana Paiva, Rui Prada, Gert Jan Hofstede, Adrie Beulens, and Ruth Aylett. 2016. Modeling culture in intelligent virtual agents. *Autonomous Agents and Multi-Agent Systems* 30 (5): 931–962. doi: <https://doi.org/10.1007/s10458-015-9312-6>.
- Mead, Ross, Amin Atrash, and Maja J. Matarić. 2013. Automated Proxemic Feature Extraction and Behavior Recognition: Applications in Human-Robot Interaction. *International Journal of Social Robotics* 5 (3): 367–378. doi: <https://doi.org/10.1007/s12369-013-0189-8>.
- Miller, Christopher. 2004. Human-computer etiquette: Managing expectations with intentional agents. *Communications of the ACM* 47 (4): 30–34. doi: <https://doi.org/10.1145/975817.975840>.
- Miller, Christopher, Peggy Wu, H. Funk, P. Wilson, and W. L. Johnson. 2006. A Computational Approach to Etiquette and Politeness: Initial Test Cases. In *Proceedings of 2006 BRIMS Conference, Baltimore, MD*.
- Mills, Sara. 2003. *Gender and politeness*. Cambridge u.a.: Cambridge University Press. Sheffield Hallam University.; England;
- Minsky, Marvin. 1977. Frame-system theory. In *Thinking: Readings in cognitive science*, eds. P. N. Johnson-Laird and P. C. Wason, 355–376. Oxford: Cambridge University Press.

- Mittelstraß, Jürgen. 2018. Wissen. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 8: Th-Z*, 2nd edn, ed. Jürgen Mittelstraß, 533-535. Stuttgart: Verlag J.B. Metzler.
- Mori, Masahiro. 2012. The Uncanny Valley: First English translation of the original essay from 1970 authorized by Mori. *IEEE Spectrum*.
- Müller-Funk, Wolfgang. 2010. *Kulturtheorie: Einführung in Schlüsseltexte der Kulturwissenschaften*, 2nd edn. Tübingen: Francke.
- Mussakhojayeva, Saida, Madi Zhanbyrtayev, Yerlik Agzhanov, and Anara Sandygulova. 2016. Who should robots adapt to within a multi-party interaction in a public space? In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 483-484. Christchurch, New Zealand: IEEE Press. <https://doi.org/10.1109/HRI.2016.7451817>.
- Nähr-Wagener, Sebastian. 2020. Anerkennungs- und Verdinglichungsprozesse im Kontext eines vergrupperten, personalisierten Webs und soziosensitiver Mensch-Technik-Interaktionen. In *Industrie 4.0, Kultur 2.0 und die Neuen Medien - Realitäten, Tendenzen, Mythen*, eds. Julius Erdmann, Björn Egbert and Sonja Ruda, 77-90.
- Nähr-Wagener, Sebastian. 2022. Sozial angemessenes Handeln-Können als situations(in)variante Kulturtechnik des Umgangs. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer.
- Neches, R., R. E. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, and W. R. Swartout. 1991. Enabling Technology for Knowledge Sharing 12 (3): 36-56.
- Newton, Casey. 2019. The secret lives of Facebook moderators in America. *The Verge*, February 25.
- Nishida, Hiroko. 2005. Cultural Schema Theory. In *Theorizing about Intercultural Communication*, ed. W. B. Gudykunst, 401-419. Thousand Oaks, London, New Delhi: Sage.
- Opp, Karl-Dieter. 1983. *Die Entstehung sozialer Normen: Ein Integrationsversuch soziologischer, sozialpsychologischer und ökonomischer Erklärungen*. Tübingen: Mohr.
- Pariser, Eli. 2017. *Filter Bubble: Wie wir im Internet entmündigt werden*, 2nd edn. München: Hanser.
- Parsons, Talcott, and Bryan S. Turner. 2005. *The social system*, 2nd edn. Abingdon, Oxon: Routledge.
- Parsons, Talcott, Edward A. Shils, and Edward Shils (eds.). 1951. *Toward a General Theory of Action*. Cambridge, Mass.: Harvard University Press.
- Pentland, Alex. 2007. Social Signal Processing [Exploratory DSP]. *IEEE Signal Processing Magazine* 24 (4): 108-111. doi: <https://doi.org/10.1109/MSP.2007.4286569>.
- Piaget, Jean. 1936. *La naissance de l'intelligence chez l'enfant*. Paris: Delachaux & Niestlé.
- Piaget, Jean. 1970. Piaget's Theory. In *Carmichael's manual of child psychology*, 3rd edn, eds. Paul Henry Mussen and Leonard Carmichael, 703-732. New York, NY: Wiley.
- Piaget, Jean. 1976. *Die Äquilibration der kognitiven Strukturen*, 1st edn. Stuttgart: Klett.
- Piaget, Jean. 2002 [1928]. *Judgement and reasoning in the child*. London: Routledge.
- Plessner, Helmuth. 2002. *Grenzen der Gemeinschaft: Eine Kritik des sozialen Radikalismus*. Frankfurt am Main: Suhrkamp.
- Poljanšek, Tom. 2022. Angemessen ist, was als angemessen passiert. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer.

- Posner, Eric A. 2000. *Law and social norms*. Cambridge, Mass.: Harvard Univ. Press.
- Pratto, Felicia, and Jim Sidanius. 2012. *Social Dominance: An intergroup theory of social hierarchy and oppression*. Cambridge: Cambridge University Press.
- Pratto, Felicia, Jim Sidanius, Lisa M. Stallworth, and Bertram F. Malle. 1994. Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of personality and social psychology* 67 (4): 741–763. doi: <https://doi.org/10.1037/0022-3514.67.4.741>.
- Pratto, Felicia, James H. Liu, Shana Levin, Jim Sidanius, Margaret Shih, Hagit Bachrach, and Peter Hegarty. 2016. Social Dominance Orientation and the Legitimization of Inequality Across Cultures. *Journal of Cross-Cultural Psychology* 31 (3): 369–409. doi: <https://doi.org/10.1177/0022022100031003005>.
- Preyer, Gerhard. 2012. *Rolle, Status, Erwartungen und soziale Gruppe*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Proyas, Alex. 2004. *I, Robot*. Overbrook Entertainment.
- Przepiorka, Wojtek, and Andreas Diekmann. 2013. Temporal Embeddedness and Signals of Trustworthiness: Experimental Tests of a Game Theoretic Model in the United Kingdom, Russia, and Switzerland. *European Sociological Review* 29 (5): 1010–1023. doi: <https://doi.org/10.1093/esr/jcs078>.
- Quaquebeke, Niels van, and Tilman Eckloff. 2010. Defining Respectful Leadership: What It Is, How It Can Be Measured, and Another Glimpse at What It Is Related to. *Journal of Business Ethics* 91 (3): 343–358. doi: <https://doi.org/10.1007/s10551-009-0087-z>.
- Quaquebeke, Niels van, Daniel C. Henrich, and Tilman Eckloff. 2007. “It’s not tolerance I’m asking for, it’s respect!”: A conceptual framework to differentiate between tolerance, acceptance and (two types of) respect. *Gruppendynamik und Organisationsberatung* 38:185–200.
- Rahwan, Iyad, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia Breazeal, Jacob W. Crandall, Nicholas A. Christakis, Iain D. Couzin, Matthew O. Jackson, Nicholas R. Jennings, Ece Kamar, Isabel M. Kloumann, Hugo Larochelle, David Lazer, Richard McElreath, Alan Mislove, David C. Parkes, Alex ‘Sandy’ Pentland, Margaret E. Roberts, Azim Shariff, Joshua B. Tenenbaum, and Michael Wellman. 2019. Machine behaviour. *Nature* 568 (7753): 477–486. <https://doi.org/10.1038/s41586-019-1138-y>.
- Rao, Anand S., and Michael Georgeff. 1995. BDI Agents: From Theory to Practice: AAAI. *Proceedings of the First International Conference on Multiagent Systems*.
- Reckwitz, Andreas. 2019. *Das Ende der Illusionen: Politik, Ökonomie und Kultur in der Spätmoderne*.
- Reeves, Byron, and Clifford Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places*. s.l.: CSLI Publications.
- Richter, Phillipp, and Andreas Kaminski. 2016. „Before you even know ...“ - Big Data und die Erkennbarkeit des Selbst. *IRIE - International Review of Information Ethics* 24 (5): 44–54.
- Rios-Martinez, J., A. Spalanzani, and C. Laugier. 2015. From Proxemics Theory to Socially-Aware Navigation: A Survey. *International Journal of Social Robotics* 7 (2): 137–153. doi: <https://doi.org/10.1007/s12369-014-0251-1>.
- Rogers, Everett M., William B. Hart, and Yoshitaka Miike. 2002. Edward T. Hall and the History of Cultural Communication: The United States and Japan. *Keio Communication Review* 24:3–26.

- Romele, Alberto. 2020. *Digital hermeneutics: Philosophical investigations in new media and technologies*. New York: Routledge.
- Romele, Alberto, Marta Severo, and Paolo Furia. 2018. Digital hermeneutics: from interpreting with machines to interpretational machines. *AI & SOCIETY* 35:73-68. doi: <https://doi.org/10.1007/s00146-018-0856-2>.
- Rumelhart, David. 1980. Schemata: The Building Blocks of Cognition. In *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education*, eds. Rand Spiro, Bertram Bruce and William Brewer, 33–58. Hillsdale: Lawrence Erlbaum Associates.
- Rumelhart, David, and Donald Norman. 1978. Accretion, Tuning, and Restructuring: Three Modes of Learning. In *Semantic factors in cognition*, eds. John Cotton and Roberta Klatzky, 37–53. Hillsdale: Lawrence Erlbaum.
- Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. 1977. The Preference for Self-Correction in the Organization of Repair in Conversation. *Language* 53 (2): 361–382.
- Salem, Maha, and Kerstin Dautenhahn. 2017. Social Signal Processing in Social Robotics. In *Social signal processing*, eds. Alessandro Vinciarelli, Judee K. Burgoon, Maja Pantic and Nadia Magnenat-Thalmann, 317–328. Cambridge: Cambridge University Press.
- Schank, Roger C. 1975. *Conceptual information processing*. New York: Elsevier.
- Schank, Roger C. 1982. *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge: Cambridge University Press.
- Schank, Roger C. 1999. *Dynamic Memory Revisited*. Cambridge: Cambridge University Press.
- Schank, Roger C., and Robert Abelson. 1977. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Schiappa, Edward, Peter B. Gregg, and Dean E. Hewes. 2005. The Parasocial Contact Hypothesis. *Communication Monographs* 72 (1): 92–115. doi: <https://doi.org/10.1080/0363775052000342544>.
- Schroll-Machl, Sylvia. 2016. *Die Deutschen - wir Deutsche: Fremdwahrnehmung und Selbstsicht im Berufsleben*, 5th edn. Göttingen: Vandenhoeck & Ruprecht.
- Schwartz, Oscar. 2019. In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation. <https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation>. Accessed 17 January 2021.
- Schwartz, Shalom H. 1977. Normative Influences on altruism. *Advances in experimental social psychology* 10:221–279. doi: [https://doi.org/10.1016/S0065-2601\(08\)60358-5](https://doi.org/10.1016/S0065-2601(08)60358-5).
- Schwartz, Shalom H. 2006. A theory of cultural value orientations: Explications and applications. *Comparative Sociology* 5 (2-3): 137–182.
- Schwartz, Shalom H. 2012. An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture* 2 (1). <https://doi.org/10.9707/2307-0919.1116>.
- Schwartz, Shalom H., and Judith A. Howard. 1981. A Normative Decision-Making Model of Altruism. In *Altruism and Helping Behavior: Social, Personality, and Developmental Perspectives*, eds. P. J. Rushton and R. M. Sorrentino, 189–211. Hillsdale: Lawrence Erlbaum.

- Schwemmer, Oswald. 2010. Kultur. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 4: Ins-Loc*, 2nd edn, ed. Jürgen Mittelstraß, 405–408. Stuttgart: J.B. Metzler.
- Schwemmer, Oswald. 2018. Wert (moralisch). In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 8: Th-Z*, 2nd edn, ed. Jürgen Mittelstraß, 462–463. Stuttgart: Verlag J.B. Metzler.
- Searle, John R. 1995. *The Construction of Social Reality*. New York: The Free Press.
- Searle, John R. 2002. *Consciousness and Language*. Cambridge: Cambridge University Press.
- Seiche, Matthias. 2018. Wert (ökonomisch). In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 8: Th-Z*, 2nd edn, ed. Jürgen Mittelstraß, 463–465. Stuttgart: Verlag J.B. Metzler.
- Si, Mei, Stacy C. Marsella, and David v. Pynadath. 2006. Thespian: Modeling Socially Normative Behavior in a Decision-Theoretic Framework. In *Intelligent virtual agents: 6th International Working Conference, IVA 2006: Marina del Rey, CA, USA, August 21-23, 2006: proceedings*, ed. Jonathan Gratch, 369–382. Berlin: Springer.
- Siep, Ludwig. 2014. *Anerkennung als Prinzip der praktischen Philosophie: Untersuchungen zu Hegels Jenaer Philosophie des Geistes*. Hamburg: Meiner.
- Siep, Ludwig. 2022. Angemessenheit und Anerkennung aus philosophischer und philosophiehistorischer Perspektive. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer.
- Solomon, Steven, Michael van Lent, Mark Core, Paul Carpenter, and Milton Rosenberg. 2008. A Language for Modeling Cultural Norms, Biases and Stereotypes for Human Behavior Models. *BRIMS*.
- Spence, Michael. 1973. Job Market Signaling. *The Quarterly Journal of Economics* 87 (3): 355–374. doi: <https://doi.org/10.2307/1882010>.
- Stahl, Titus. 2014. Anerkennung, Subjektivität und Gesellschaftskritik. *Deutsche Zeitschrift für Philosophie* 62 (2): 239–259. doi: <https://doi.org/10.1515/dzph-2014-0019>.
- Standage, Tom. 2002. *Der Türke. Die Geschichte des ersten Schachautomaten und seiner abenteuerlichen Reise um die Welt*. Frankfurt/Main: Campus-Verl.
- Stemmler, Gerhard, Dieter Bartussek, Dirk Hagemann, Frank Spinath, Manfred Amelang, Marcus Hasselhorn, Silvia Schneider, and Wilfried Kunde. 2016. *Differentielle Psychologie und Persönlichkeitsforschung*, 8th edn. Stuttgart: W. Kohlhammer Verlag.
- Tajfel, H., and J. C. Turner. 1986. The Social Identity Theory of Intergroup Behavior. In *Psychology of Intergroup Relation*, eds. S. Worchel and W. G. Austin, 7–24. Chicago: Hall Publishers.
- Taylor, Charles. 2009. *Multikulturalismus und die Politik der Anerkennung*. Frankfurt am Main: Suhrkamp-Verl.
- Taylor, Shelley E., and Jennifer Crocker. 1981. Schematic bases of social information processing. In *Social cognition*, eds. E. T. Higgins, C. P. Herman and M. P. Zanna, 89–134. Hillsdale, NJ: Erlbaum.
- Tennyson, Matthew F., Deitra A. Kuester, John Casteel, and Christos Nikolopoulos. 2016. Accessible Robots for Improving Social Skills of Individuals with Autism. *Journal of Artificial Intelligence and Soft Computing Research* 6 (4): 164. doi: <https://doi.org/10.1515/jaiscr-2016-0020>.

- Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudge: Wie man kluge Entscheidungen anstößt*, 3rd edn. Berlin: Econ.
- Thomas, Alexander. 2016. *Interkulturelle Psychologie: Verstehen und Handeln in internationalen Kontexten*. Göttingen: Hogrefe (Therapeutische Praxis).
- Thorndyke, Perry. 1984. Applications of schema theory in cognitive research. In *Tutorials in learning and memory*, eds. Stephen M. Kosslyn and John R. Anderson, 167–192. San Francisco: W. H. Freeman.
- Ting-Toomey, Stella. 1999. *Communicating across cultures*. New York: The Guilford Press.
- Tolmach Lakoff, Robin, and Mary Bucholtz (eds.). 2004. *Language and woman's place: Text and commentaries*. New York: Oxford University Press.
- Torta, Elena, Raymond H. Cuijpers, and James F. Juola. 2013. Design of a Parametric Model of Personal Space for Robotic Social Navigation. *International Journal of Social Robotics* 5 (3): 357–365. doi: <https://doi.org/10.1007/s12369-013-0188-9>.
- Trompenaars, Fons, and Charles Hampden-Turner. 1993. *Riding the Waves of Culture: Understanding Cultural Diversity in Business*. London: Nicholas Brealey.
- TU Dortmund. 2020a. MoFFa. <https://ts.sowi.tu-dortmund.de/forschung/projekte/moffa-abgeschlossen/>. Accessed 7 December 2020a.
- TU Dortmund. 2020b. Regelungssystemtechnik | Fakultät für Elektrotechnik und Informationstechnik. Accessed 7 December 2020b.
- Tulving, Endel. 1985. How many memory systems are there? *American Psychologist* 40 (4): 385–398. doi: <https://doi.org/10.1037/0003-066X.40.4.385>.
- Turkle, Sherry. 2011. *Alone together: Why we expect more from technology and less from each other*. New York, NY: Basic Books.
- Turner, Roy M. 1994. *Adaptive reasoning for real-world problems: A schema-based approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Uexküll, Jacob von. 1920. *Theoretische Biologie*. Berlin: Paetel.
- Ullmann-Margalit, Edna. 1977. *The Emergence of Norms*. Oxford: Clarendon Press.
- Useem, John, and Ruth Useem. 1967. The Interfaces of a Binational Third Culture: A Study of the American Community in India. *Journal of Social Issues* 23 (1): 130–143. doi: <https://doi.org/10.1111/j.1540-4560.1967.tb00567.x>.
- Vallor, Shannon. 2015. Moral Deskillling and Upskilling in a New Machine Age: Reflections on the Ambiguous Future of Character. *Philos. Technol.* 28 (1): 107–124. doi: <https://doi.org/10.1007/s13347-014-0156-9>.
- Vannucci, Manila. 2014. The spatial frequencies influence the aesthetic judgment of buildings transculturally. *COGNITIVE NEUROSCIENCE* 5 (3-4): 143–149.
- Vester, Heinz-Günter. 2009. *Kompendium der Soziologie I: Grundbegriffe*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden.
- Vilhjalmsson, Hannes, W. Lewis Johnson, Peggy Wu, Christopher Miller, and Harry Funk. 2007. A Computational Approach to Etiquette and Politeness: An “Etiquette Engine™” for Cultural Interaction Training. *Proceedings of the 2007 Conference on Behavior Representation In Modeling and Simulation*.
- Vincent, James. 2018. Google Assistant adds Pretty Please functionality, updated lists, and other features. <https://www.theverge.com/2018/11/29/18117585/google-assistant-pretty-please-thank-you-polite-feature-lists-notes-stories-holiday-updates>.

- Vinciarelli, Alessandro, Maja Pantic, and Hervé Bourlard. 2009. Social Signal Processing. Survey of an emerging domain. *Image and Vision Computing* 27 (12): 1743–1759. <https://doi.org/10.1016/j.imavis.2008.11.007>.
- Vinciarelli, Alessandro, Maja Pantic, Dirk Heylen, Catherine Pelachaud, Isabella Poggi, Francesca D’Errico, and Marc Schroeder. 2012. Bridging the Gap between Social Animal and Unsocial Machine: A Survey of Social Signal Processing. *IEEE Transactions on Affective Computing* 3 (1): 69–87. doi: <https://doi.org/10.1109/T-AFFC.2011.27>.
- Vinciarelli, Alessandro, Judee K. Burgoon, Maja Pantic, and Nadia Magnenat-Thalmann (eds.). 2017. *Social signal processing*. Cambridge: Cambridge University Press.
- Visser, Thomas, Martijn Vastenburger, and David V. Keyson. 2011. Designing to Support Social Connectedness: The Case of SnowGlobe 5:129–142.
- Vogt, Catharina. 2022. Respekt als Basis sozialer Angemessenheit. In *Soziale Angemessenheit - Forschung zu Kulturtechniken des Verhaltens*, eds. Jacqueline Bellon, Bruno Gransche and Sebastian Nähr-Wagener. Wiesbaden: Springer.
- Voss, Thomas. 2001. Game theoretical perspectives on the emergence of social norms. In *Social norms*, eds. Michael Hechter and Karl-Dieter Opp, 105–138. New York N.Y.: Russell Sage Foundation.
- Waldenfels, Bernhard (ed.). 1998. *Der Anspruch des Anderen: Perspektiven phänomenologischer Ethik*. München: Fink.
- Watts, Richard J. 2005. Linguistic politeness research. *Quo vadis? In Politeness in language: Studies in its history, theory and practice*, 2nd edn, eds. Richard J. Watts, Ide Sachiko and Konrad Ehlich, xi–xlvi. Berlin u.a.: Mouton de Gruyter.
- Weber, Max. 1978. *Economy and Society. An Outline of Interpretive Sociology: Edited by Guenther Roth and Claus Wittich*. Berkeley, Los Angeles, London: University of California Press.
- Weber, J. Mark, Shirli Kopelman, and David M. Messick. 2004. A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and social psychology review: An official journal of the Society for Personality and Social Psychology, Inc* 8 (3): 281–307. https://doi.org/10.1207/s15327957pspr0803_4.
- Whitworth, Brian. 2005. Polite computing. *Behaviour & Information Technology* 24 (5): 353–363. doi: <https://doi.org/10.1080/01449290512331333700>.
- Wiegerling, Klaus. 2016. Umfassende IT-Systeme. In *Handbuch Medien- und Informationsethik*, ed. Jessica Heesen, 217–226. Stuttgart: J.B. Metzler.
- Wilson, David Sloan. 2007. *Evolution for everyone: How Darwin’s theory can change the way we think about our lives*. New York: Delacorte Press.
- Wimmer, Reiner. 2005. Macht. In *Enzyklopädie Philosophie und Wissenschaftstheorie: Bd. 5: Log-N*, 2nd edn, ed. Jürgen Mittelstraß, 164–167. Stuttgart: J.B. Metzler.
- Wu, Peggy, Christopher A. Miller, Harry Funk, and Vanessa Vikili. 2011. Computational Models Of Etiquette And Culture. In *Human-computer etiquette: Cultural expectations and the design implications they place on computers and technology*, eds. Caroline C. Hayes and Christopher Allan Miller, 63–90. Boca Raton, Fla.: CRC Press.
- Wulf, Christoph. 2022. Soziale Angemessenheit durch mimetische, rituelle und repetitive Prozesse. In *Soziale Angemessenheit – Forschung zu Kulturtechniken des Verhaltens*, eds. Bellon et al., Wiesbaden: Springer.
- Wyer, Robert S. 2012. A Theory of Social Information Processing. In *Theories of social psychology*, ed. Paul A. van Lange, 156–177. Los Angeles, Calif.: Sage.

- Yamamoto, Michiya, and Tomio Watanabe. 2006. Time Lag Effects of Utterance to Communicative Actions on CG Character-Human Greeting Interaction. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2006: Sept. 6–8: Piscataway, NJ*, 629–634.
- Zahavi, Amotz. 1975. Mate selection - A selection for a handicap. *Journal of theoretical biology* 53 (1): 205–214.
- Zhang, Tao, Biwen Zhu, and David D. Kaber. 2011. Anthropomorphism and Social Robots: Setting Etiquette Expectations. In *Human-computer etiquette. Cultural expectations and the design implications they place on computers and technology*, eds. Caroline C. Hayes and Christopher Allan Miller, 231–262. Boca Raton, Fla.: CRC Press.
- Ziem, Alexander. 2008. *Frames und sprachliches Wissen: Kognitive Aspekte der semantischen Kompetenz*. Berlin: Walter de Gruyter.
- Zola-Morgan, Stuart, and Larry Squire. 1990. Neurophysiological investigations of memory and amnesia: Findings from humans and non-human primates. In *The development and neural bases of higher cognitive functions*, ed. Adele Diamond, 732–740. New York, NY: New York Academy of Sciences.