



# The FASA model

# 4

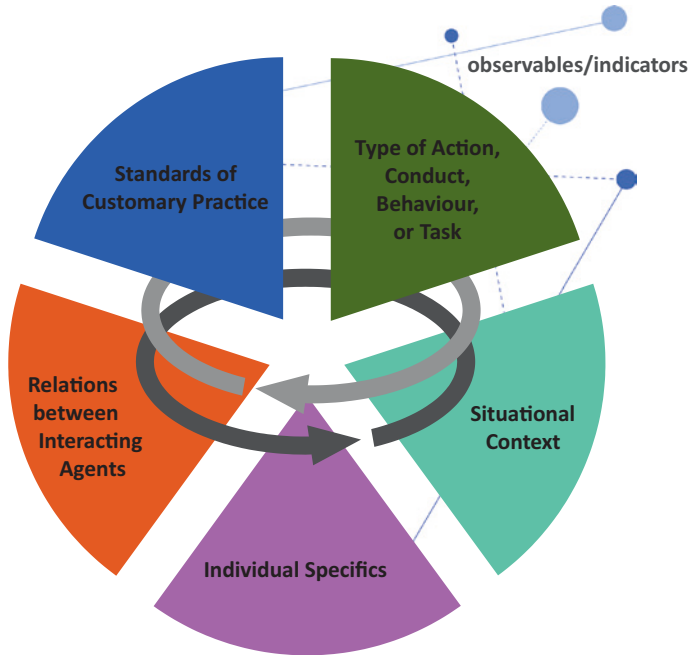
We don't all carry an etiquette handbook, but everyone seems to know good manners. Although violations exist, most people are polite most of the time. The level of conformity is striking, a fact blurred by vivid memory for occasional lapses. Politeness is ubiquitous, and it's practised automatically. 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.<sup>1</sup> 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»,<sup>2</sup> e) «Standards of Customary Practice».** In the presence of complex interdependencies—**represented by the circular**

---

<sup>1</sup>These 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).

<sup>2</sup>Here, 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 evaluation<sup>3</sup> of appropriateness include: «Standards of

<sup>3</sup> '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 observable<sup>4</sup> ‘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

---

<sup>4</sup>Simplified 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)

<b>«Type of Action, Conduct, Behaviour, or Task»</b>		
<p><b>«Time»</b> 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><sup>a</sup>): 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><b>«Role identities»</b> 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)

<b>«Type of Action, Conduct, Behaviour, or Task»</b>		
<b>«Intention»</b> 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?
<b>«Consequences»</b> 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»)?

<sup>a</sup>See 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 organisation<sup>5</sup> names a total of eight “**primary social interaction schemas**” (PSI SCHEMAS, cf. Nishida 2005, p. 407):

<b>Fact-and-concept schemas</b>	for example “Tokyo is the capital of Japan” or “Bicycles are those vehicles that have two wheels, a seat, and handlebars” (ibid.)
<b>Person schemas</b>	for example “John is neurotic”, “Taro is shy”, “Mary is easy-going” (ibid.)
<b>Self schemas</b>	for example <i>I am generous</i> or <i>Other people perceive me as generous</i> .
<b>Role schemas</b>	contain social stereotypes such as <i>Women work hard</i> , <i>Men cook well</i> , or <i>Poor people have good manners</i>
<b>Context schemas</b>	contain <b>rules of appropriateness</b> , 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>
<b>Procedure schemas (scripts)</b>	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>

---

<sup>5</sup> ‘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*.<sup>6</sup>

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,

---

<sup>6</sup>For 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).<sup>7</sup>

---

<sup>7</sup>According 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.<sup>8</sup>

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

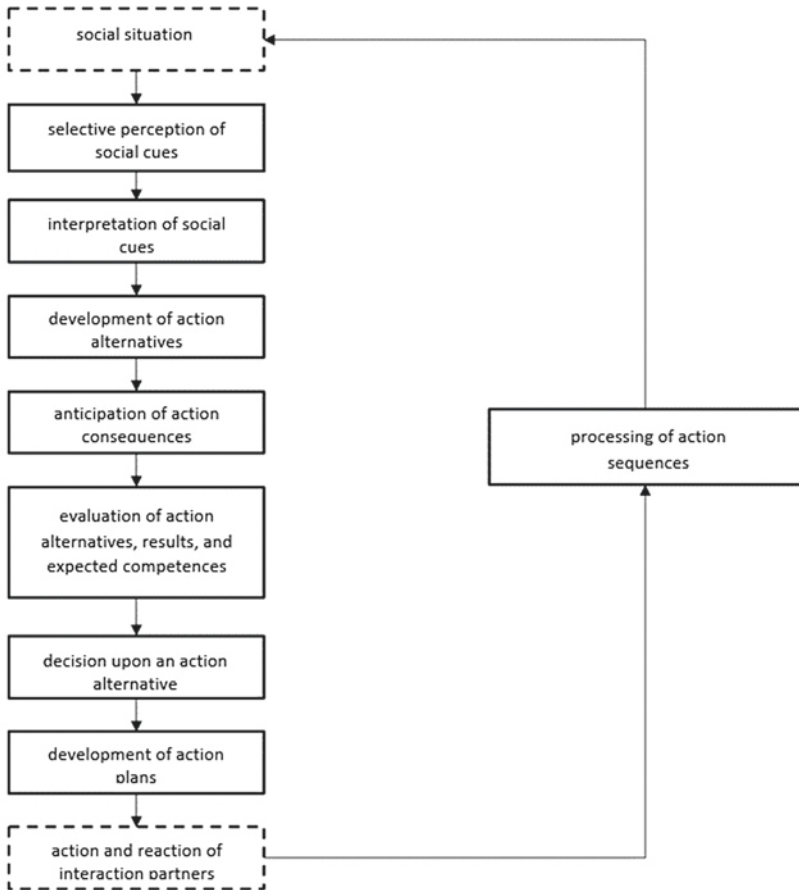
- (a) relevant signs from the social situation are not perceived or are misinterpreted;
- (b) possible alternative courses of action are not taken into consideration;
- (c) consequences of action are ignored or evaluated incorrectly;
- (d) alternative courses of action are associated with unrealistic 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).

<sup>8</sup>For 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.



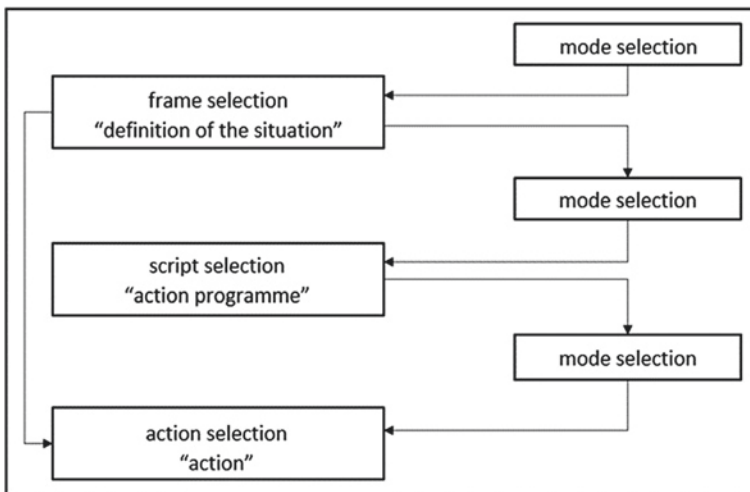
If this information is not perceived, the interaction partner will not adapt their behaviour to accommodate the other person's haste and may therefore hold them up inappropriately instead of showing polite consideration. Mistakes can also be made when interpreting social cues. The non-hurrying person 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]).<sup>9</sup> 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.

---

<sup>9</sup>This 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*<sup>10</sup> (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

---

<sup>10</sup>NUDGE 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)

<b>&lt;&lt;Situational Context&gt;&gt;</b>		
<b>&lt;Place and time&gt;</b> Where and when is the behaviour taking place?	<b>&lt;Place&gt;</b> : 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?
	<b>&lt;Time&gt;</b> : 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?
<b>&lt;Framing&gt;</b> ‘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’?	
	<b>&lt;Media-based and performative mediation&gt;</b> 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.<sup>11</sup>

- 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.

---

<sup>11</sup>That 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).

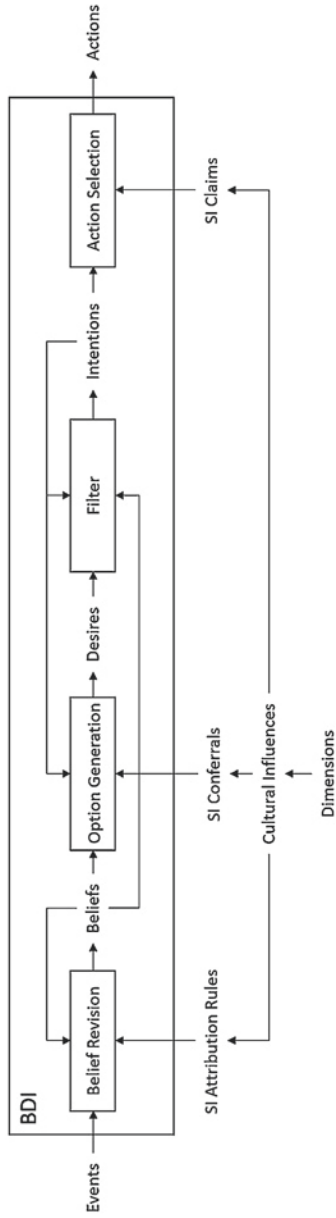


Mascarenhas et al. identify three levels that are important for the development of sociocultural agents and regarding situations: an **interaction** level, a **group** level, and a **society** level (Mascarenhas et al. 2016, 937ff.).

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»<sup>12</sup> 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.).

---

<sup>12</sup> 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.<sup>13</sup>

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).

---

<sup>13</sup>However, 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)

<b>«Individual Specifics»</b>	
<b>«Personal evaluation structures»</b> 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?
<b>«Personal characteristics»</b> 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?
<b>«Individual shaping of social roles»</b> 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**.

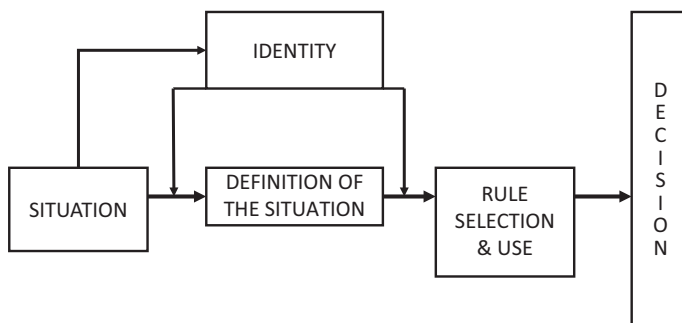
The definition of the situation informs the person about the norms, expectations, rules, learned behaviours, skills, and possible strategies that are relevant. (...) The definition of the situation suggests a choice set. Choosing among the options, we contend, is a rule-directed exercise. (ibid. p. 285)

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).

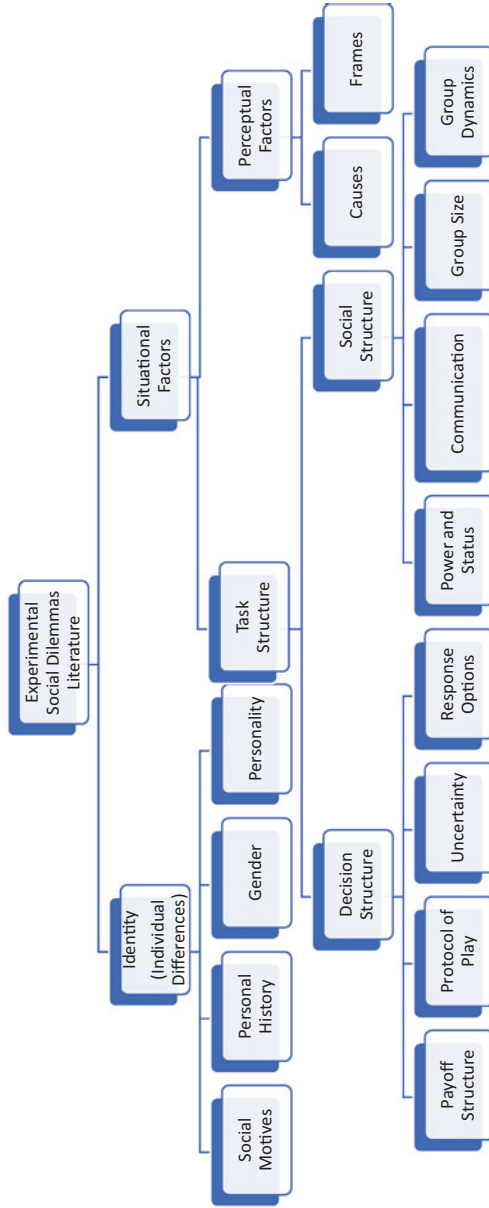
Finally, an analysis of the literature on decisions in social dilemmas (cf. Figure 4.6) provides Weber, Kopelman, and Messick with variables that offer more precise insight into the **decisive factors of the subjective definition of the situation** (ibid. p. 287):

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)<sup>14</sup>. 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

---

<sup>14</sup>Own 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)

<b>«Relations between Interacting Agents»</b>	
<b>«Familiarity and relationship aspects»</b> 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)
<b>«Intention»</b> 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**.<sup>15</sup> 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.

---

<sup>15</sup> 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**’.<sup>16</sup> 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 party<sup>17</sup> (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’);

---

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

<sup>17</sup>Which 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)

<b>«Standards of Customary Practice»</b>	
<b>«Habitus»</b> 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?
<b>«Values/social norms»</b> 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.)
<b>«Regulative norms»</b> 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.<sup>18</sup> 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).<sup>19</sup> This can take the form of imperatives and prohibitions, duties and responsibilities

---

<sup>18</sup>Regarding 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.).

<sup>19</sup>For 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.<sup>20</sup> 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'.<sup>21</sup> 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.

---

<sup>20</sup>This 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.

<sup>21</sup>Any '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 theory<sup>22</sup>:

#### 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

---

<sup>22</sup>In 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.<sup>23</sup>

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.<sup>24</sup> 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

---

<sup>23</sup> 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).

<sup>24</sup> 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.

