

This is a section of [doi:10.7551/mitpress/11252.001.0001](https://doi.org/10.7551/mitpress/11252.001.0001)

# The Handbook of Rationality

Edited by: Markus Knauff, Wolfgang Spohn

## Citation:

*The Handbook of Rationality*

Edited by: Markus Knauff, Wolfgang Spohn

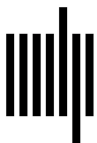
DOI: 10.7551/mitpress/11252.001.0001

ISBN (electronic): 9780262366175

Publisher: The MIT Press

Published: 2021

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.



The MIT Press

## 10.3 Rationality in Communication

Georg Meggle

### Summary

*Communicative actions are actions aiming at being understood.* This is the common ground of many different theories of communication. Explicating what is entailed in this smallest common denominator is the first and foremost task of a general theory of communicative actions. This chapter covers three different kinds of communicative theory: an individualistic theory of communication following ideas of Paul Grice (section 1) and two collectivist theories of communication (section 2), namely, speech act theory, initiated by John L. Austin and John Searle, and the theory of communicative actions as developed by Jürgen Habermas. Finally, in section 3, we examine various aspects of rationality involved in and presupposed by these different approaches.

### 1. The Individualistic Approach

This chapter is not on communication in general, but only about communicative actions. A general theory of such communicative actions is thus part of a general theory of actions. Its main question is, What is the *differentia specifica* of communicative actions? And regarding this handbook, is the rationality of communicative actions really a form of rationality *sui generis* (as vigorously argued by Habermas)?

*Communicative actions are actions aiming at being understood, and understanding an action is just knowledge of its meaning.* This is the common ground of all sorts of communication theories. Thus, explicating what is entailed by this common ground is the foremost task of the theory sought after. However, these explications divide over the readings of “action,” “action understanding,” and “action meaning” they are starting from.

Regarding actions, one has to distinguish between action types (e.g., the action type *singing*) and action tokens (e.g., my singing now). Theories whose basic concepts are primarily defined for action tokens, as performed by their respective agents or subjects at particular

times, follow the so-called research program of *methodological individualism*. Theories following the alternative program of *methodological collectivism* start from concepts basically related to action types. Notice that methodological differences need not entail ontological differences. It is this methodological difference that sets the stage for the following overview.<sup>1</sup>

### 1.1 Basic Concepts

*Communicative actions are actions aiming at being understood.* This is enough for them to qualify as instrumental actions, that is, as actions by means of which the agent tries to achieve her respective aim(s).

Instrumental or intentional actions can be defined quite easily by combining the three basic (qualitative) elements of any (individualistic) action theory:

$D(X, f) := X$  does or performs (at time  $t$ ) the action  $f$ ;

$B(X, A) := X$  believes (strongly)—i.e., is convinced—(at  $t$ ) that  $A$ ;

$W(X, A) := X$  wants or desires (at  $t$ ) that  $A$ .

These are the three basic notions on which everything else will build. First the notion of intentional action:

*Definition 1:* By doing  $f$ ,  $X$  intends to achieve aim  $A$  iff (if and only if)  $X$  does  $f$ ,  $X$  wants that  $A$ , and  $X$  believes that  $A$  will come about if  $X$  does  $f$ ; in symbols:

$I(X, f, A) := D(X, f) \wedge W(X, A) \wedge B(X, D(X, f) \leftrightarrow A)$ .

Here,  $\leftrightarrow$  stands for material equivalence.<sup>2</sup> Here and in the following, it should be strengthened to a subjunctive biconditional. However, we better neglect this complication.

Such an action is *successful* iff  $X$ 's belief, her expectation of success, is correct, that is, iff  $D(X, f) \leftrightarrow A$  indeed holds. Sometimes our aims have a certain hierarchy. If in  $X$ 's view, her aim  $A_2$  is merely a means to achieve aim  $A_1$ , aim  $A_1$  is her *primary aim* regarding  $A_2$ . Therefore, *the primary aim*, if it exists, is the aim that is primary regarding all other aims.

What, then, does it mean to understand an action? It is common to say that we understand an action of  $X$  iff we know the reasons for which it was done. These reasons should be spelled out more generally with reference to  $X$ 's beliefs and desires, or preferences and probabilities (see chapter 2.1 by Broome and chapter 8.2 by Peterson, both in this handbook). However, since our focus is on intentions (definition 1), we may conveniently restrict our considerations to the following special case: we understand an action iff we know the intention with which it was performed.

Since intentions are subjective (in the sense that having an intention is being convinced of having it), this agrees with Weber (1913/1988, pp. 429, 432), who says that we understand an action iff we know its subjective meaning. Here, the subjective meaning of an action is nothing but the meaning or function this action has for the agent. And this we know when we know the corresponding intention with which the action was performed.

Let us make our notion of understanding a bit more explicit. For the sake of simplicity, we may equate *knowledge* with true belief (conviction): " $X$  knows (at  $t$ ) that  $A$ " =  $K(X, A) := B(X, A) \wedge A$ , although we are aware that since Plato, this equation is taken not to be correct. Then our notion comes to

*Definition 2:*  $Y$  (fully) understands  $X$ 's doing  $f$  iff  $Y$  knows all the aims which  $X$  intends to achieve by doing  $f$ ; symbolically:  $U(Y, D(X, f)) := K(Y, I(X, f, A^*))$ , where  $A^*$  represents the sum of all aims  $X$  intends by doing  $f$ .

The upshot is that the understanding of actions and the rationality of actions are two sides of the same coin. Every rational action is understandable, and according to definition 2, *only* rational actions are understandable—trivially so: reasons, and intentions in particular, can be (fully) known to us only if they exist.

## 1.2 Communicative Actions

Let us apply now these basic notions to communicative actions. Typically, such actions are utterances (of sentences). However, they may also be gestures, facial expressions, flying a flag, and so on. Almost anything may serve communication. We shall focus here, *pars pro toto*, on one general form of communication attempts, namely, so-called *informatives*, where the primary communicative aim of speaker  $S$  is to get the hearer  $H$  to believe some proposition  $p$ .

So,  $S$ 's doing  $f$  is an (informative) *communicative attempt* addressed at  $H$  with the content  $p$ , only if by doing  $f$ ,  $S$  intends to make  $H$  believe (at  $t' > t$ ) that  $p$ ; in symbols:  $CA(S, H, f, p) \rightarrow I(S, f, B'(H, p))$ .<sup>3</sup>

Communicative aims may also be achieved in a non-communicative way. The crucial question hence is, What distinguishes a communicative action  $f$  from a merely (noncommunicative) instrumental action with the same aim? It is obviously the special communicative way by means of which  $S$  expects to achieve her primary communicative aim, namely, through  $H$ 's very recognition of her communication attempt  $f$ . That is,  $S$  believes that  $H$  comes to believe  $p$  (at  $t'$ ) iff  $H$  recognizes (knows at  $t'$ ) that  $S$ 's doing  $f$  is such a communicative attempt. Formally:  $CA(S, H, f, p) \rightarrow B(S, B'(H, p) \leftrightarrow K'(H, CA(S, H, f, p)))$ .

When we take the two conditions on communicative attempts just stated to be jointly sufficient, we arrive at the following adequacy condition:

$$AC1: CA(S, H, f, p) \leftrightarrow I(S, f, B'(H, p)) \wedge B(S, B'(H, p) \leftrightarrow K'(H, CA(S, H, f, p)))$$

Note that this is not a definition of CA; as such, it would be circular. Hence, it only serves as a condition that must be derivable from any acceptable definition of CA.

The most important conclusion from AC1 is the following reflexivity condition:

$$RC: CA(S, H, f, p) \rightarrow I(S, f, K'(H, CA(S, H, f, p)))$$

In terms of definition 2, RC states the crucial principle that communication always aims at being understood. It follows from AC1, if one accepts the principle that believed means and consequences of intended aims are intended as well.

The next step is to realize that this reflexivity condition in turn spawns further consequences and entails the unlimited transparency or openness of communicative intentions. That is, the communicative attempt has the primary intention  $I_0 = I(S, f, B'(H, p))$  that  $H$  should believe  $p$ . Via RC, this entails that  $I_1 = I(S, f, K'(H, I_0))$  also belongs to the intentions of the communicative attempt. Then, so does  $I_2 = I(S, f, K'(H, I_1))$ , and so on. In general, if  $I_n$  is a communicative intention, so is  $I_{n+1} = I(S, f, K'(H, I_n))$  for any  $n \geq 0$ . In this way, the reflexivity condition RC extends to

$$RC^*: CA(S, H, f, p) \rightarrow I(S, f, K'(H, I^*)),$$

where  $I^*$  represents the conjunction of the hierarchy just developed; it is, we might say, an unlimitedly transparent communicative intention.

However, we still don't know what an adequate explication of CA looks like. Reviewing the various answers proposed would be a lengthy procedure (sketched in Meggle, 1981). Let us cut the long story short and just look at its beginning and its end. It starts with Grice's

basic model (in Grice, 1957, p. 384), which says the following in our present terms:

$$CA(S, H, f, p) := I(S, f, B'(H, p)) \wedge B(S, B'(H, p)) \leftrightarrow K'(H, I_0).$$

And we have just developed the end of the story:

$$\text{Definition 3: } CA(S, H, f, p) := \\ I(S, f, B'(H, p)) \wedge B(S, B'(H, p)) \leftrightarrow K'(H, I^*).$$

The difference between  $I_0$  and  $I^*$  seems small—merely an asterisk. Yet this asterisk harbors what we have called “unlimited transparency” or “openness.” Note that definition 3 indeed entails the adequacy condition AC1, since, according to this definition, knowledge of the communicative attempt  $CA(S, H, f, p)$  amounts to knowledge of the unlimited intention  $I^*$ .

In definition 2, we explicated understanding an action in general. This applies as well to communicative actions and results in

$$\text{Definition 4: Person } Y \text{ (who may be the hearer } H \\ \text{ himself) understands } S\text{'s doing } f \text{ as a communicative} \\ \text{attempt addressed at } H \text{ with the content } p \text{ iff } Y \text{ knows} \\ \text{that } S \text{ undertakes such a communicative attempt. In} \\ \text{symbols: } U(Y, CA(S, H, f, p)) := K(Y, CA(S, H, f, p)).$$

The definition takes this simple form because knowing of the communicative attempt  $CA(S, H, f, p)$  entails, via the strong reflexivity condition RC\*, knowing of the speaker's unlimited intention  $I^*$ , with which she performs this attempt.

Let's still attend to the hearer's side. Suppose that you understand  $S$ 's action as a communicative attempt addressed at you with the content  $p$ . When would you actually correspond to  $S$ 's involved expectation of success, that is, actually believe  $p$ ? Normally only if you take both conditions,

$$\text{Sincerity: } CA(S, H, f, p) \rightarrow B(S, p), \text{ and}$$

$$\text{Correctness: } B(S, p) \rightarrow p,$$

to be satisfied, that is, that  $S$  does not tell a lie and does not err. I refer to these conditions as “communicative normality conditions.” They may or may not be actually satisfied in a particular given case. On the individual level on which we have moved so far, a theory of communicative actions may be silent on such normality conditions. However, we will see that they play a crucial role on the social level.

### 1.3 Intersubjective Meaning

If the individualistic approach is to be successful, the social level needs to be construed on its basis. Explaining how this might work in principle is the goal of this

section. A first noteworthy point is an ontological one: whereas subjective meanings (i.e., intentions) are those of a communicative action *token* as performed by a subject  $S$  (at a time  $t$ ), intersubjective meanings are attached to action *types* relative to a collective or a population  $P$  and, moreover, relative to a certain type of situation  $\Sigma$ . The meaning of a certain gesture or utterance may vary from group to group and even from situation to situation within one group. Again, we better speak of action types generally and not specifically of utterance types.

Now, what might intersubjective meanings be? How to explicate  $M(P, \Sigma, F, p)$ : that action type  $F$  has for  $P$  in  $\Sigma$  the *intersubjective meaning* that  $p$ ? It seems clear that such meanings must be common knowledge in the group  $P$ . We follow the usual definition and say that  $p$  is *common knowledge* in  $P$ —symbolically:  $K^*(P, p)$ —iff everybody in  $P$  knows that  $p$  ( $K(S, p)$  for all  $S$  in  $P$ ), everybody in  $P$  knows that everybody in  $P$  knows that  $p$ , everybody knows this in turn, and so on (see also chapter 9.2 by Perea on common belief, this handbook). Thus, intersubjective meanings should satisfy the following adequacy condition:

$$\text{AC2: } M(P, \Sigma, F, p) \rightarrow K^*(P, M(P, \Sigma, F, p)).$$

On the basis of our above treatment of the individual level, it is not difficult to arrive at an adequate explication of  $M$ . We might start with the proposal of a mere *communicative regularity* (where the index 0 at  $M_0$  signals that there is still a caveat).

$$\text{Definition 5: } M_0(P, \Sigma, F, p) := \text{for each speaker } S \text{ in } P \\ \text{and each action token } f \text{ of type } F \text{ performed in a situ-} \\ \text{ation of type } \Sigma, \text{ there is a hearer } H \text{ in } P \text{ such that} \\ \text{doing } f \text{ is a communicative attempt with the con-} \\ \text{tent } p: D(S, f) \rightarrow CA(S, H, f, p).^4$$

We might call  $M_0$  “regular meaning,” since it states only a regularity in communication tokens. The caveat is that  $M_0$  does not yet fulfill our adequacy condition AC2. But this is easily filled by requiring that regular meaning be common knowledge:

$$\text{Definition 6: } M(P, \Sigma, F, p) := K^*(P, M_0(P, \Sigma, F, p)).$$

It may be easily shown that this definition satisfies AC2; if the regular meanings are common knowledge, so are the intersubjective meanings (for all details, see Meggle, 2003, pp. 244–251).

Definitions 5 and 6 are the essential steps from (subjective) meanings on the action-*token* level to (intersubjective) meanings on the action-*type* level. Notice that the latter meanings are *defined* in terms of—but *not identified* with—the former ones. (This difference is mostly disregarded or misinterpreted in the criticisms of the

individualist approach. This misinterpretation was unfortunately invited by Bennett [1973], one of the champions of the Gricean doctrine. Already the title of this paper, “The Meaning-Nominalist Strategy,” suggested that conventional meaning is just a special case of speaker’s meaning,  $CA(S, H, f, p)$ . This is false, of course.)

An essential step in our account is still missing. On the individual level, a speaker  $S$  makes a communicative attempt  $CA(S, H, f, p)$  with two sorts of expectations, namely, the expectation that the hearer understands her attempt,  $B(S, D(S, f) \rightarrow K'(H, CA(S, H, f, p)))$ , and that her attempt is indeed successful, that is,  $B(S, K'(H, CA(S, H, f, p)) \rightarrow B'(H, p))$ . It is not obvious how these expectations transfer to the social level. On which reasons can they be based in order to make them well founded on the action-type level?

Concerning the expectation of understanding, the answer is just  $M(P, \Sigma, F, p)$  itself and its entailed common knowledge; each speaker expects her audience to understand her. However, concerning the expectation of success, there is obviously a rationality gap left open by definition 6. Up to now, there is nothing in  $M(P, \Sigma, F, p)$  on which this expectation could regularly keep relying. But without such a reliable base, the communicative regularity of  $M(P, \Sigma, F, p)$  would be miraculous and could not be stable.

The missing link adding stability to regularity is *conventionality*, which is based not only on common knowledge but also on common interests. For the *locus classicus* of a game-theoretical modeling of how such interests or preferences turn regularities into conventions, see Lewis (1969); see also chapter 9.1 by Albert and Kliemt (this handbook). Lewis’s explication of signaling conventions is an explication of conventional meaning as a special case of  $M(P, \Sigma, F, p)$  (see Skyrms, 2010, or the reconstructions in von Kutschera, 1983, and Meggle, 2010).

The way conventionality closes the rationality gap behind  $M(P, \Sigma, F, p)$  is best explained by an idealized story, in which  $\Sigma = \Sigma^*$  is a *perfect communication situation*, which means (i) it is a common interest of  $S$  and  $H$  in  $\Sigma^*$  that  $B'(H, p) \leftrightarrow p$ ; (ii) it is common knowledge among  $S$  and  $H$  that  $K(S, p)$  or  $K(S, \neg p)$ , but neither  $K(H, p)$  nor  $K(H, \neg p)$ ; and (iii) it is also common knowledge among  $S$  and  $H$  that  $S$  and  $H$  will follow a joint strategy consisting of  $S$ ’s strategy  $D(S, f) \leftrightarrow p$  and of  $H$ ’s strategy  $B'(H, p) \leftrightarrow D(S, f)$ .

Now, suppose that  $S$  in  $P$  makes the communicative attempt  $CA(S, H, f, p)$  toward  $H$  in  $P$ , in such a perfect situation  $\Sigma^*$ , that  $f$  is of type  $F$  and that indeed the intersubjective meaning of  $F$  in  $\Sigma^*$  is  $p$ , that is,  $M(P, \Sigma, F, p)$ . Then the normality conditions of sincerity and correctness (section 1.2) do hold. This is entailed by clause (iii) about

common knowledge of the joint strategy, while clauses (i) and (ii) explain why  $S$  even makes the effort of a communicative attempt. In this way, the best reasons for the expectation of success at the *CA-token* level (i.e., the conditions of sincerity and correctness) are transformed also into the best reasons at the *CA-type* level.

Clearly, perfect communication situations are very idealized. The actual picture is statistically blurred everywhere. However, the ideal picture is the guideline for all attempts at more realism. Here, the point was only to explain the individualistic strategy for construing intersubjective meanings in terms of subjective communicative actions.

#### 1.4 Toward the Semantics and Pragmatics of Expressions

So far, the individualistic approach has delivered a meaning theory for action performances (e.g., whistling, singing, or flying a flag). However, these performances must be distinguished from their products or results (whistles, songs, or the flags flied). Thus, we get the following matrix comprising the four different levels:

	Action performance	Action product
Type	Action type (II)	Product form (III)
Token	Action token (I)	Product token (IV)

We have dealt with I and II. So, the remaining big question is, How do we get from an action-theoretic semantics for communicative actions to a semantics for action products (signs or expressions)? Let’s turn first to level III and then to level IV.

For simple (unstructured) signs (like green or red lights, flags, etc.) the solution is simple, too: their meanings can be identified with the conventional meanings of their performances. So, if  $F$ -doing is the way of producing the sign- or utterance-type  $U$ , we may equate  $M(P, \Sigma, U, p) = M(P, \Sigma, F, p)$ . However, this trivial identification in the case of unstructured signs is not available when structured signs enter the field. Then the question arises as to how the meaning of a complex sign is a function of the meanings of its parts. And as every linguist will agree, there is no generally accepted answer to this question up to now.

As for structured expressions—linguistic expressions being their most important case—a truly systematic action-theoretic semantics would have to start from an exclusively syntactically characterized language (say, predicate logic amended by all kinds of indexicals and intensional and performative operators) and deliver an interpretation of this language, as paradigmatically designed in modern

intensional semantics. And it would have to show how this interpretation can be systematically based on corresponding communicative conventions stated at level II (a task already solved by Lewis [1969], extended by von Kutschera [1983], and embedded into a full account of level I by Meggle [2010]). Eventually, the entire feat would have to be carried over to natural languages. Thereby, however, we would get deeply involved in linguistics.

Let us finally turn to level IV, accounting for the communicative meanings of concrete utterances  $u$  of type  $U$ . Since a semantics for utterance types  $U$  (level III) is already presupposed here, level IV is the field of traditional pragmatics. Thus, the main question on this level is how best to systematize the various literal and nonliteral ways of doing things with signs and words in concrete situations (in a communicative form).

Although pragmatics is the most prolific field in communication theory and has shown already tremendous progress with regards to formalizations in the past decades (see, e.g., Horn & Ward, 2004), a systematic pragmatics based on, and formulated in terms of, a general communication theory is still missing. For quite obvious reasons: conventional sign meanings in the group  $P$  are relative to various (idealized) situation-types  $\Sigma$ . Thereby, literal sign-usages in a specific situation  $\sigma$  of type  $\Sigma$  are associated with mutual communicative expectations firmly grounded in the conventions behind  $M(P, \Sigma, F, p)$ ; how this works was indicated at the end of section 1.3 for those perfect situations. However, what the mutual communicative expectations actually are in the situation  $\sigma$  is an entirely open question, because  $\sigma$  has many other features beyond being of type  $\Sigma$ . Each feature is potentially relevant; this might involve everything commonly known from human social life.

Well, relevant for *what* exactly? (See Wilson & Sperber [2012] and chapter 4.3 by Merin, this handbook.) This crucial question is best answered by returning to the beginning of our individualistic account, where we tried to capture communicative understanding.

We started with the question of what communicative understanding means in the case of a totally unspecified action token, and now we have arrived at the same question for concrete utterances of action types with meanings already specified in relation to particular populations, situations, languages, and other structured instruments of human interaction (argumentation rules and other cultural rules and habits in the fields of knowledge, sciences, religion, jurisdiction, arts, etc.). However, all these kinds of factors are relevant only insofar as they contribute to the simple question explicitly left open at the beginning (but required to get a good answer at this final level): how

can speaker  $S$  addressing hearer  $H$  in this specific situation  $\sigma$  reasonably believe her expectation of understanding and of success to be fulfilled? Well, all these factors are relevant insofar as they are part and parcel of the best rational explanation of  $S$ 's communicative attempt.

There are, from ancient rhetorics (see chapter 5.6 by Woods, this handbook) over speech act theory, Habermas's universal pragmatics (both to be discussed in a moment), Grice's theory of implicatures (Grice, 1989), and the just-mentioned relevance theory up to modern linguistic pragmatics, many proposals to structure this extremely broad field of possibly relevant factors by various sorts of conversational rules or maxims (see, e.g., Korta & Perry, 2006/2020). However, all of this is beyond the scope of this chapter.

## 2. Collectivistic Approaches

### 2.1 Speech Act Theory

Whereas an individualist theory of communicative actions starts with actions, as well as their meaning and understanding, in the subjective sense, speech act theory—which is perhaps the leading collectivist version following the meaning-as-use approach of Wittgenstein (1953)—proceeds from actions already endowed with an intersubjective meaning. In terms of the matrix above, speech act theory starts at level II and not at level I.

Accordingly, the basic question of speech act theory is not how to define communicative actions in general but how to account for the different forms of these acts as specified in the following open list of so-called illocutionary roles: commanding, asking, questioning, thanking, threatening, warning, recommending, dissuading, declaring, excusing, promising, and so on. This has been the central concern of speech act theory as founded by Austin (1962) and developed by Searle (1969) (see also Vanderveken, 1990–1991). Unfortunately, though, no explicit definition has been presented going beyond such an open list of illocutionary roles or acts having these roles. So, on the one hand, speech act theory was very fruitful in many applications in philosophy and linguistics, but on the other hand, it was also in need of conceptual explication from the start.

In fact, it is not quite correct to say that speech act theory starts at level II. Although it focused on illocutionary roles and acts, it presupposed that these acts are utterances of signs already having a conventional linguistic meaning.<sup>5</sup> Thus, it was exclusively interested in linguistic communication. In other words, it started rather at a combination of levels II and III, where the latter was assumed not to be in need of further clarification.

This assumption has initially barred speech act theory from developing a general pragmatic semantics that explicates meaning also for structured (linguistic) action products. The first to notice and try to overcome this bar was Alston (1964; and in more detail in Alston, 2000); he was followed by von Savigny (1983) and in quite different ways by Brandom (1994).

However, let me rather dwell on how speech act theory construes the relation between levels I and II. It is here where its differences with the individualistic approach in section 1 become most conspicuous. A central distinction maintained in all versions of speech act theory is that between illocutionary and perlocutionary acts, the latter ones being acts of actually bringing about something by intentionally doing or uttering something, with or without the help of conventions. And this fits exactly our explication of instrumental actions as presented in definition 1. In short: intentional perlocutionary acts *are* instrumental acts.

Now, the central claim of speech act theory associated with this distinction is the *irreducibility thesis*: illocutionary acts are not definable in terms of perlocutionary acts, and illocutionary forces are not definable in terms of perlocutionary effects; both are *sui generis*. (The locus classicus here is Austin, 1962, chapter 9.)

Is the irreducibility thesis correct? This depends on what kind of illocutionary acts we are talking about. Following earlier suggestions of Strawson (1964) and Schiffer (1972), it is mainly due to Bach and Harnish (1979) that we should distinguish between two kinds of illocutionary acts: they are either communicative acts or institutional acts. Paradigm cases of the latter would be the (procedural) act types of getting married, baptizing, or betting at an auction—the very cases Austin had started from.

Now, by performing such an institutional act, one may indeed try to achieve various things (say, enrichment, in-group acceptance, or eternal redemption). The point, though, is that no special perlocutionary aspect, no special instrumental intention, is conceptually required for an act to count as institutional. Thus, for this kind of illocutionary acts, the irreducibility thesis holds trivially. However, this does not entail that it holds also for the other sort of illocutionary acts, namely, the communicative ones. On the contrary, as shown in section 1, the thesis is simply false at least for the informatives discussed there. There we found that some (communicative) illocutionary intentions *are* a special sort of perlocutionary (instrumentalist) intentions. Hence, at least as far as the informatives, an embrative class of communicative illocutionary acts, are concerned, speech act theory could be individualistically reconstructed on the

basis of the explications given in section 1, as shown by Schiffer (1972) and Meggle (1981).

The most basic issue between an individualist and a collectivist approach to meaning revolves around what kind of explication of  $M(P, \Sigma, F, p)$  one is willing to accept. Whereas the individualist would accept a reductionist explication as given in section 1, this very sort of reductionism is denied by the collectivist. He would agree to the adequacy condition AC2 for intersubjective meaning stated in section 1.3. However, he cannot accept the individualistic explication of  $M$  as presented in definitions 5 and 6 (section 1.3). He is committed to give  $M$  and its corresponding notion of understanding a different reading, one that refers to the relevant rules defining what is or counts as a *correct performance* of the respective action type  $F$  in  $P$  and  $\Sigma$ . His notion of understanding is simply this: we understand a communicative action iff we know its rules.<sup>6</sup>

## 2.2 Habermas's Theory of Communicative Actions

*Communicative actions*, as this term is used in Habermas (1981), are of a very different sort from all those considered so far. Habermas conceives of them as collective actions, performed by a collective  $P$  and not by a single individual, as assumed in the individualistic approach (section 1) and in speech act theory (section 2.1). If such individual actions of members of  $P$  are to add up to a communicative action in Habermas's sense, they must integrate into some kind of common practice in which the members of  $P$  engage with a certain shared attitude. It is this shared attitude that characterizes Habermasian communicative action for a collective  $P$  in a situation  $\Sigma$  with respect to some state of affairs  $p$ . This attitude is best described as embracing three connected parts: (i) a cognitive, (ii) a volutative, and (iii) a praxiological part.

Ad (i): it is common knowledge in  $P_\Sigma$  (i.e., the subgroup of  $P$  involved in  $\Sigma$ ) that it is (maybe only tentatively) an open question for  $P_\Sigma$  whether  $p$  or not- $p$ , where  $p$  refers to a factual (empirical) or expressive (referring to one's state of mind) or normative state of affairs. Ad (ii): it is a common interest in  $P_\Sigma$  to engage jointly and open-mindedly in trying to reach mutual understanding—in the sense of mutual agreement (consensus) in  $P_\Sigma$ —about whether  $p$  or not- $p$  is the case. And ad (iii): the procedures to achieve this common goal are those of giving and taking reasons in consideration of  $p$ . They are governed by the essential characteristics of our everyday speech acts, and their rules agree with, and could be explicated in terms of, the argumentation rules of rational theoretical or practical discourse (see also chapter 5.5 by Hahn & Collins and chapter 5.6 by Woods, both in this handbook).

As suggested by part (iii), this characterization of Habermasian communicative actions utilizes, as he intends, some strongly idealized aspects of speech act theory. Following Bühler's (1934) tripartition of the organum model of speech events (with its three poles of sender, receptor, and things and, respectively, its relations of expression, appeal, and representation), Habermas's speech act model distinguishes between three worlds (reflected in part (i) above), with which every speech act is correlated: the objective, the social, and the subjective world, which correspond to the distinction between culture, society, and personality. According to Habermas's theory, these three dimensions show up in three corresponding dimensions of communicative rationality. Communication aims at the discursive redemption of validity claims in ways structurally and universally embedded in our everyday speech act practice. This aim is at the center of Habermas's attempts at a unified concept of rationality. These attempts were strongly assisted by his friend K.-O. Apel. Apel's work on transcendental pragmatics<sup>7</sup> incorporated Peirce's pragmatic conception of truth as consensus reached in the long run by an indefinitely extended community of discourse and thus set the stage for the speech-act-theoretic backings of Habermas's theory.

As should be clear from this very condensed summary, Habermas's project of spelling out the consequences of the above basic characterization of communicative actions in his sense, and applying them to the various possible social and political fields, is much richer than a specific theory such as the one sketched in section 1. It covers a whole research program, which we cannot even sketch here. Moreover, one has to keep in mind how Habermas develops and presents his program, namely, in his typical multiperspective style by means of systematic and historical reconstructions and by combining philosophical accounts with empirical approaches from psychology and the social and political sciences. Commonly, such research programs can only be meant as work in progress from the outset.

Already this very brief outline shows that switching from the notion of communicative actions as developed in section 1 to Habermas's notion implies a radical paradigm shift. The shift may be required, since Habermas intended his theory to be a starting point for a new foundation of social theory, particularly of a theory of modern societies and their transnational extensions. Since this paradigm shift has often been disregarded (maybe even by Habermas himself—see the last paragraph of this section), let us at least notice some of its aspects.

For this purpose, we should keep in mind that we distinguished above between (a) instrumental acts, (b)

communicative attempts in general, and (c) conventional and linguistic communicative attempts in particular and, in section 2.1, between (d) communicative and (e) institutional illocutionary acts. Thus, if one would want to systematically reconstruct Habermas's theory, one would have to ask how it relates to these distinctions or, more generally, how Habermas's collective communicative actions are correlated with the various communicative and illocutionary acts on the individual or on the group level. Let me only make a few suggestions in this direction.

For instance, Habermas's theory is committed to the irreducibility thesis (section 2.1). Hence, one would have to know how its speech-act-theoretic foundations look like without this thesis, in order to fairly compare his theory with the individualistic approach at the general level (b).

Next, according to Habermas, communication is essentially a consensus-oriented activity. This famous claim can be subscribed to by a Gricean individualist, too. However, this orientation toward consensus would have to be spelled out quite differently for the two sides. For ease of comparison, let us restrict the group  $P$  to  $S$  and  $H$ . Then, according to the above sketch of Habermas's approach,  $S$  and  $H$  try to reach common knowledge about the truth of either  $p$  or not- $p$ . In the individualistic approach, by contrast, a communicative attempt of the form  $CA(S, H, f, p)$  only entails that  $S$  intends by doing  $f$  that she will bring about common knowledge between her and  $H$  that  $CA(S, H, f, p)$ . So, on both sides, there is necessarily some openness involved. In Habermas's account, however, it is a joint open-mindedness related to some unknown factual or normative state of affairs  $p$  of common interest, while in the individualistic approach, this openness is essentially restricted to the communicative state of affairs  $CA(S, H, f, p)$  itself.

Moreover, for Habermas, communicative actions are opposed to strategic actions, trivially so, because he defines strategic actions with reference to their missing openness. Therefore, Habermas seems to imply that strategic actions are also the opposite of communicative actions in the individualistic sense. However, there is no such opposition. As explained in section 1, communicative actions in the latter sense are instrumental actions; they need not even be sincere, although they have to be so sufficiently often on the conventional level.

Finally, what about Habermas's central tenet that communicative actions are actions *sui generis*? That is, how do Habermasian communicative actions relate to the various actions types (a)–(e)? There is an easy first answer confirming this tenet: since collective actions in general transcend the contribution of all individual actions, they



are by definition *sui generis*. And so are, hence, Habermas's communicative actions. However, nothing follows from this with respect to whether or not even these communicative actions may be ultimately explicated in terms of more basic individualistic concepts.

### 3. Rationality Assumptions—An Overview

#### 3.1 Rationality in the Individualistic Approach

Let us look back on our route and see what kinds of rationality assumptions we have used at the various stages. And let us first take up the individualistic approach (section 1). There we first came across considerations of rationality when we explained a general concept of understanding an action as knowing its reasons. These reasons lie in the agent's preferences and beliefs, the yardstick by which her action's rationality is measured. This is the concept of *action rationality* (as dealt with in chapter 2.1 by Broome and chapter 2.2 by Wedgwood and the entire section 8 in this handbook).

Our explanation assumed that these reasons in turn correspond to certain standards of rationality that they have to satisfy in order to qualify as reasons in the first place. Which standards? They certainly include the requirement that both, the beliefs and the desires in question, are consistent. One might impose the requirement of deductive closure:

$$B(X,A) \wedge B(X,A \rightarrow B) \rightarrow B(X,B).$$

One might extend this to the desires. In any case, such principles concern the rationality of our *attitudes*. This is not the place to discuss them in detail, but see, for example, chapter 5.1 by van Ditmarsch, chapter 8.1 by Grüne-Yanoff, and chapter 11.1 by Horty and Roy (all in this handbook). (My own compass has been Lenzen, 1980.) Of course, they would have to include principles stating rationality conditions among the various attitudes, such as the following bridge principle:

$$W(X,A) \wedge B(X,A \rightarrow B) \rightarrow W(X,B).^8$$

Such principles, which mediate among various attitudes held by a person at one time, need to be complemented by principles that mediate between these attitudes at different times. Again, this is not the place to expand on these issues (see chapter 5.2 by Rott; chapter 5.3 by Kern-Isberner, Skovgaard-Olsen, & Spohn; chapter 5.4 by Gazzo Castañeda & Knauff; and section 5 of chapter 8.1 by Grüne-Yanoff, all in this handbook). Finally, the fact that the agent has sufficient reasons for a given action, and that those reasons in turn satisfy the relevant rationality requirements, does not mean she will

actually do it. She will only do it if she herself *is* rational in the given situation. So, in a nutshell, we must distinguish three aspects of rationality already within general action: action rationality, rationality of an action's reasons, and personal rationality.

Owing to the thoroughly subjective nature of action rationality, the question of whether the agent's attitudes themselves are well founded in turn remains open, especially whether her action (as Weber, 1913/1988, puts it) is rational not only in a subjective means–end sense but also in an objective-correctness sense (by being based on true beliefs). The latter is required only for a successful intentional action.

What do communicative actions add to all this? To wit, nothing! Except certain more special expectations on the part of the speaker. The speaker expects the hearer to understand her communication attempt and thus to assume of her all three forms of rationality. Moreover, as a result of the reflexivity of communication, she expects that the hearer also realizes that the speaker expects him to make these rationality assumptions of her, and so forth. Hence, *communicative rationality* is nothing but a special case of general action rationality plus its openness intended by the speaker.<sup>9</sup> The soundness of the reasons of a communicative action pertains only to its success. However, within the framework of a general theory, one can initially avoid saying anything about whether the expectations of understanding and success have to be sound themselves.

The picture changes as soon as the *types* of communicative actions receive an intersubjective meaning. If *S* and *H* belong to the relevant population, the expectation of understanding is supported by the relevant meaningful situations. Hence, a communicative attempt is not only subjectively rational regarding its aim of being understood but also objectively rational in such a situation. Moreover, the expectation of communicative success must be confirmed with sufficient frequency. Otherwise, the relevant communicative regularity would sooner or later collapse. Communication by means of signs or expressions with a regular, conventional, or linguistic meaning differs from communication without this backup precisely by the regular confirmation of the communicative aims. The structure of this confirmation can be explicated—following the ideas of Lewis (1969)—by applying game theory as a collective version of rational decision theory (see chapter 9.1 by Albert & Kliemt, chapter 9.2 by Perea, chapter 9.3 by Alexander, chapter 10.2 by Schmid, and chapter 10.4 by Raub, all in this handbook). However, note that “collective” is to be understood here only *in sensu diviso*; the agents are, as in decision making in general, individual persons, members of a collective, but not the collective itself.

With this subjective and intersubjective backing of the expectations of understanding, further communicative spaces open up, especially for “communication between the lines,” as studied, for example, in Grice’s theory of conversational implicatures, which is based on his “cooperative principle.” This is the fundamental principle of his theory but a derived principle of cooperative rationality (see Grice, 1989).

### 3.2 Rationality in Collectivistic Approaches

Let us turn to how communicative rationality may be conceived within speech act theory. If speech act theory were robbed of the irreducibility thesis (section 2.1), then it would coincide with the individualistic theory. Or so we have argued at least with respect to informatives. Hence, the relevant rationality assumptions would coincide as well.

But what about the collectivist version of speech act theory including the irreducibility thesis? In this case, the central concept of understanding, with which our search for rationality assumptions involved in communication started, does not refer to the intention with which the relevant communicative action was performed. It rather relates to the rules with reference to which a correct performance of the relevant action type is defined. Now, just as one might say that rationality is what makes an action correct in a certain sense, one might say that it is the constitutive rules for illocutions that make them correct according to speech act theory. Compare, for example, if you want to play chess (where it would be redundant to add “correctly”), you have to stick to the chess rules. But this would not give us a rational explanation why you would like to play chess right now or why you prefer to make this move rather than that. So, even in speech act theory, one would need to refer to the above notions of action rationality.

Finally, as far as Habermas’s theory of communication is concerned, I do not know of any overall analytic reconstruction, although there are some promising beginnings (see Steinhoff [2009] and Heath [2003], the former in a more destructive, the latter in a more constructive way). In particular, I do not know of any serious attempts to connect Habermas’s collectivistic account with the field of collective intentionality richly developed in the past 25 years (see the other chapters of section 10 in this handbook). That is regrettable, since communicative actions in Habermas’s sense may be expected to play a central role within this field. Without doubt, it would be very promising to develop Habermas’s program further within the framework of theories of collective intentionality.

### Notes

1. For more details, see Meggle (2010).
2. The point of the equivalence is this: requiring only the weaker belief  $B(X, D(X, f) \rightarrow A)$  would allow many actions to be believed to be sufficient for the aim  $A$ . However, in the theory of communicative actions, we would like to be able to explain why exactly action  $f$  and not any other action sufficient for  $A$  is performed. And this explanation requires the stronger belief as stated.
3. For directives, the corresponding primary communicative aim would be that  $H$  performs some action  $a$ . The logical structure of directives and informatives is the same: just substitute the aim  $D'(H, a)$  for the aim  $B'(H, p)$ .
4. This definition, indeed the entire section, is replete with universal quantifiers. Of course, this is only an ideal type. Actually, all these quantifiers should be statistically weakened to high percentages, as has been suggested by Lewis (1975).
5. Cf. Austin’s explanation of what he called a “locutionary act” in Austin (1962, pp. 92–93): it is an utterance with a linguistic meaning “in the sense of Frege.”
6. For the different sorts of rules constitutive for communicative meanings in language games versus rules of chess, see Meggle (1985) and, in the same spirit, Kemmerling (1992).
7. See Apel (1976, 1981) and Kuhlmann (1992).
8. For instance, Kant in *Groundwork of the Metaphysics of Morals* (1785/1903, B 44–45) states, “He who wants the purpose (given that the reason has a decisive impact on his actions) also wants the means necessary for the purpose.” The quote suggests that the material implication within the belief operator needs to be strengthened to a subjunctive conditional or a strict implication.
9. That is, the speaker expects common knowledge of rationality, as studied in epistemic game theory (see chapter 9.2 by Perea, this handbook).

### References

- Alston, W. (1964). *Philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall.
- Alston, W. (2000). *Illocutionary acts and sentence meaning*. Ithaca, NY: Cornell University Press.
- Apel, K.-O. (1976). *Transformation der Philosophie* [Transformation of philosophy] (Vol. I). Frankfurt/Main, Germany: Suhrkamp.
- Apel, K.-O. (1981). *Transformation der Philosophie* [Transformation of philosophy] (Vol. II). Frankfurt/Main, Germany: Suhrkamp.
- Austin, J. L. (1962). *How to do things with words*. Oxford, England: Clarendon Press.
- Bach, K., & Harnish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.

- Bennett, J. (1973). The meaning-nominalist strategy. *Foundations of Language*, 10, 141–168.
- Brandom, R. B. (1994). *Making it explicit*. Cambridge, MA: Harvard University Press.
- Bühler, K. (1934). *Sprachtheorie* [Theory of language]. Jena, Germany: Fischer.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377–388.
- Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Habermas, J. (1981). *Theorie des kommunikativen Handelns* [Theory of communicative action] (2 vols.). Frankfurt/Main, Germany: Suhrkamp.
- Heath, J. (2003). *Communicative action and rational choice*. Cambridge, MA: MIT Press.
- Horn, L. R., & Ward, G. (Eds.). (2004). *The handbook of pragmatics*. Oxford, England: Blackwell.
- Kant, I. (1903). *Grundlegung zur Metaphysik der Sitten* [Groundwork of the metaphysics of morals]. In I. Kant, *Kant's gesammelte Schriften* (Vol. 4, pp. 385–463). Berlin, Germany: Reimer. (Original work published 1785)
- Kemmerling, A. (1992). Bedeutung und der Zweck der Sprache [Meaning and purpose of language]. In W. Vossenkuhl (Ed.), *Von Wittgenstein lernen* (pp. 99–120). Berlin, Germany: Akademie-Verlag.
- Korta, K., & Perry, J. (2020). Pragmatics. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <https://plato.stanford.edu/archives/spr2020/entries/pragmatics/>
- Kuhlmann, W. (1992). *Sprachphilosophie – Hermeneutik – Ethik: Studien zur Transzendentalpragmatik* [Philosophy of language – hermeneutics – ethics: Studies in transcendental pragmatics]. Würzburg, Germany: Königshausen und Neumann.
- Lenzen, W. (1980). *Glauben, Wissen und Wahrscheinlichkeit* [Belief, knowledge, and probability]. Vienna, Austria: Springer.
- Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Lewis, D. (1975). Languages and language. In K. Gunderson (Ed.), *Minnesota studies in the philosophy of science* (Vol. VII, pp. 3–35). Minneapolis: University of Minnesota Press.
- Meggle, G. (1981). *Grundbegriffe der Kommunikation* [Basic concepts of communication]. Berlin, Germany: De Gruyter.
- Meggle, G. (1985). Wittgenstein – Ein Instrumentalist? [Wittgenstein – an instrumentalist?] In D. Bimbacher & A. Burkhardt (Eds.), *Sprachspiel und Methode* (pp. 71–88). Berlin, Germany: De Gruyter.
- Meggle, G. (2003). Common belief and common knowledge. In M. Sintonen, P. Ylikoski, & K. Miller (Eds.), *Realism in action* (pp. 244–251). Dordrecht, Netherlands: Kluwer.
- Meggle, G. (2010). *Handlungstheoretische Semantik* [Action-theoretical semantics]. Berlin, Germany: De Gruyter.
- Schiffer, S. R. (1972). *Meaning*. Oxford, England: Clarendon Press.
- Searle, J. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, England: Cambridge University Press.
- Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford, England: Oxford University Press.
- Steinhoff, U. (2009). *The philosophy of Jürgen Habermas: A critical introduction* (K. Schönner, Trans.). Oxford, England: Oxford University Press.
- Strawson, P. F. (1964). Intention and convention in speech acts. *Philosophical Review*, 73, 439–460.
- Vanderveken, D. (1990–1991). *Meaning and speech acts* (2 vols.). Cambridge, England: Cambridge University Press.
- von Kutschera, F. (1983). Remarks on action-theoretical semantics. *Theoretical Linguistics*, 10(1), 1–11.
- von Savigny, E. (1983). *Der Begriff der Sprache* [The concept of language]. Stuttgart, Germany: Reclam.
- Weber, M. (1988). Über einige Kategorien der verstehenden Soziologie [On some categories of understanding sociology]. In J. Winkelmann (Ed.), *Gesammelte Aufsätze zur Wissenschaftslehre* (pp. 427–474). Tübingen, Germany: J. C. B. Mohr (Paul Siebeck). (Original work published 1913)
- Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge, England: Cambridge University Press.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, England: Blackwell.

© 2021 The Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Knauff, Markus, editor. | Spohn, Wolfgang, editor.

Title: The handbook of rationality / edited by Markus Knauff and Wolfgang Spohn.

Description: Cambridge : The MIT Press, 2021. | Includes bibliographical references and index.

Identifiers: LCCN 2020048455 | ISBN 9780262045070 (hardcover)

Subjects: LCSH: Reasoning (Psychology) | Reason. | Cognitive psychology. | Logic. | Philosophy of mind.

Classification: LCC BF442 .H36 2021 | DDC 153.4/3—dc23

LC record available at <https://lcn.loc.gov/2020048455>