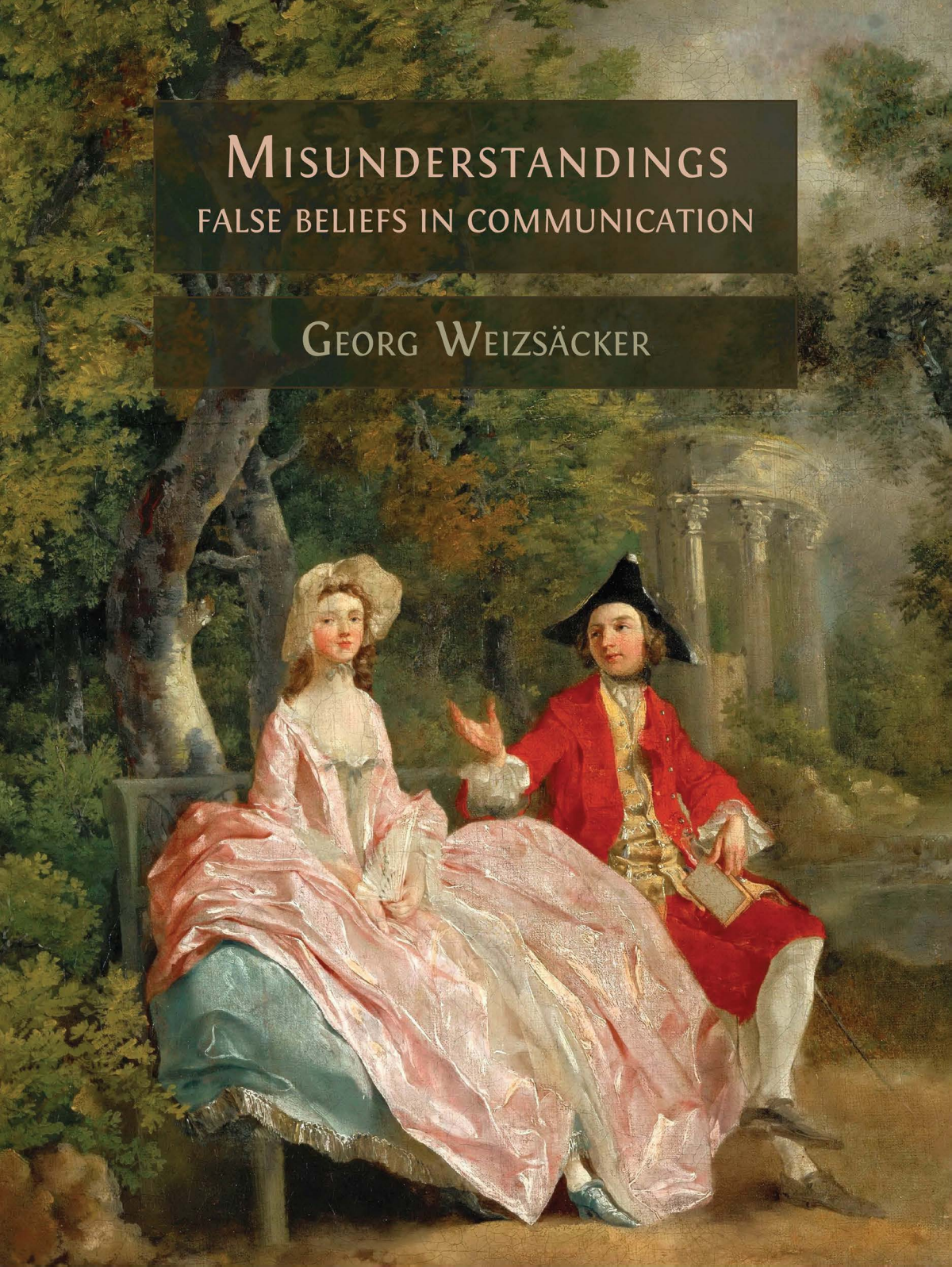


# MISUNDERSTANDINGS FALSE BELIEFS IN COMMUNICATION

GEORG WEIZSÄCKER



<https://www.openbookpublishers.com/>

©2023 Georg Weizsäcker

This work is licensed under an Attribution-ShareAlike 4.0 International (CC BY-SA 4.0). This license allows users to share, copy and redistribute the material in any medium or format; to adapt, remix, transform, and build upon the material for any purpose, even commercially, and only so long as attribution is given to the creator. Attribution should include the following information:

Georg Weizsäcker, Misunderstandings: False Beliefs in Communication.  
Cambridge, UK: Open Book Publishers, 2023,  
<https://doi.org/10.11647/OBP.0367>

In order to access detailed and updated information on the license, please visit  
<https://doi.org/10.11647/OBP.0367#copyright>

Further details about the CC BY-SA 4.0 license are available at  
<https://creativecommons.org/licenses/by-sa/4.0/>

All external links were active at the time of publication unless otherwise stated and have been archived via the Internet Archive Wayback Machine at  
<https://archive.org/web>

Digital material and resources associated with this volume are available at  
<https://doi.org/10.11647/OBP.0367#resources>

ISBN Paperback: 978-1-80064-881-4  
ISBN Hardback: 978-1-78374-051-2  
ISBN Digital (PDF): 978-1-80511-138-2  
DOI: 10.11647/OBP.0367

Cover: Thomas Gainsborough, Conversation in a Park (1746), Louvre Museum,  
[https://commons.wikimedia.org/wiki/File:Thomas\\_Gainsborough\\_-\\_Conversation\\_in\\_a\\_Park\\_\(1746\).jpg](https://commons.wikimedia.org/wiki/File:Thomas_Gainsborough_-_Conversation_in_a_Park_(1746).jpg)  
Cover design: Jeevanjot Kaur Nagpal

The publication of this work was supported by the Open Access Publication Fund of Humboldt-Universität zu Berlin, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 491192747. The publication of this work was further supported by the CRC/Transregio 'Rationality and Competition: The Economic Performance of Individuals and Firms' (TRR 190), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 280092119.



# Chapter 2

## Having a conversation

### Three conversations

Agniezka and Dimitri stand in silence as the elevator moves up one floor, and another one. Dimitri then addresses the issue of the lost competition: “This was bloody. Not the end of the story, though. We will hit back.”

Agniezka turns her head and gives Dimitri an expressionless stare. She says, “Hopeless. Make it your bloodbath, not mine. I’m done with this project, sorry.” The door opens and Agnieszka leaves the elevator.

Rachel straightens her back. “Mr. Governor, please consider the urgency of research on stolen artifacts. We ask you to look into the possibility of funding a new center of competence that studies the provenance of cultural property. How did this property leave its home? The Cambodian statues in western museums and collections are a good example. Many of them were looted in postwar turmoil. This was publicized in newspapers and we verified the reports with scientific studies. But for contexts outside of Cambodia, the facts about provenance are mostly unresearched. We need more structured knowledge, we have enormous work before us and we need your help with it.”

The governor hears her out and replies, “Rachel, another great line of research! We are on the same page and I much admire your spirit and that of your colleagues. I well recall how you and I first met, in a panel discussion on cultural policies, planning a common agenda. Let’s continue this process! You also know that we already do a lot for top-level research in our museums, through our funding programs.” He pauses briefly. “We will continue these efforts, too – but here and now, we have to move on.” He looks at the



researcher sitting next to Rachel.

For a few seconds Steve looks at Ralph, the older boy, who turns around and looks back at Steve. His shoulders are not drooping any more.

“Wanna play?”, says Steve, holding the ball in his hands.

Ralph’s face assumes its usual, hostile look. He walks straight towards Steve and slowly, almost routinely, pushes him with his overweight arms, then walks away.

## Simplifying restrictions

A conversation has multi-faceted content, much more than meets the eye. The three conversations of Rachel, Dimitri and Steve are similar only at the surface – the talker wants the listener to do something, the listener rejects it – but they strongly differ in terms of their contexts, their human characters, their languages, their outcomes, and in the future events that they induce.

The analysis starts by narrowing things down. It focuses on what is uncertain. When people talk and listen, they do this with incomplete knowledge. Three aspects of uncertainty arise: about *actions*, *issues*, and *types*.

*Actions* are everything that people do during or after the conversation. This includes all their possible statements: statements are actions.

To notate the action (or statement) of person  $i$ , the book uses the symbol  $a^i$  and it is understood that the action is one out of a given set  $A^i$  of possible actions for person  $i$ , i.e.,  $a^i \in A^i$ .

The book will keep this and other mathematical notations to a minimum and it sometimes even describes variants of the basic set-up without notating them formally. For example, a person’s action may sometimes depend on another person’s action without a mathematical notation for it.

*Issues* are things that are exogenous to the actions. They are things that cannot be changed. The book refers to them as the “state of the world” (or simply “state”), which is notated by  $\omega$  and is, generally speaking, unknown

at the time of the conversation. The true state is an element of a larger set of possible states,  $\omega \in \Omega$ .

Not all is unknown of course, as everyone knows *something* about the state of the world. The information that person  $i$  has about  $\omega$  is notated by the symbol  $I_\omega^i$ . It is convenient to describe  $I_\omega^i$  as a subset of  $\Omega$  that contains  $\omega$ ; this indicates that person  $i$ , if she pays attention to her information, can deduce that some states are ruled out: those that lie in  $\Omega$  but not in  $I_\omega^i$ . She does not know the exact value of  $\omega$  but she can safely conclude (later, we will say, it is “manifest” to her) that  $\omega$  lies in  $I_\omega^i$ .

Most likely, the people in the conversation each know something different. The conversation may help to exchange their views.

*Types*, notated by  $\theta^i$  for person  $i$ , describe the person’s preferences. They are person-specific and include everything that is idiosyncratic: desires and tastes, capabilities and necessities.

Note that a person’s preference for certain outcomes can be related to her behavioral constraints. If person  $i$  strongly dislikes an outcome, this is similar to the case where her action set  $A^i$  does not contain an action that leads to this outcome. The book’s analysis keeps these things apart, for simplicity:  $\theta^i$  describes person  $i$ ’s preferences for a fixed set  $A^i$ , i.e., it describes which of the available possible outcomes (and actions) she likes relatively better than others.

In its general discussion, appearing in normal font and black color, the book does not specify the exact nature of actions, issues and types. In any given conversation, they may be either explicit and in the open, or hidden in the background. For a successful analysis of a given situation, one should specify these aspects in sufficient detail but also simplify where appropriate, focusing on the main features of the situation.

At the surface, Steve’s simple question (“Wanna play?”) asks about Ralph’s preference regarding a ball game. Ralph’s possible reactions, however, are not all playful: he may turn to violent bullying, especially now that Steve saw him in a state of despair. The analysis may therefore collect Steve’s possible statements into two groups: those that address Ralph’s state of despair, e.g. by asking about it, versus those that address more normal things. We also separate the set of Ralph’s possible reactions in two groups, describing whether he is violent or not, and likewise, his set of possible types:

his despair may either fuel his tendency to be violent, or it may reduce it. Steve's own type, in contrast, plays no important role in this conversation. The relevant issues, or states of the world, are the unknown events that underlie Ralph's despair. A further relevant unknown are Steve's future actions: will he tell other kids about what he saw, or not?

Rachel focuses her statement on a set of future actions that she hopes to provoke – the governor's possible support for a new research center. She could talk about many different things but here, too, we may simplify by restricting her choice to be between two topics: the process on cultural policy taht has already begun (and that the governor refers to in his response) or the new topic of provenance. We may also describe the governor's possible reactions as binary: either supporting Rachel's research with new funds, or not supporting it. We note that the relevant state of the world comprises not only the substantive topic of conversation but also the professional context, i.e., the unusual situation of the governor's visit as well as the presence of other colleagues. We also note that the personalities, or types, of Rachel and the governor are in the background; no-one addresses them openly – only the personal tone of the governor's response alludes to them.

Dimitri is hoping to keep Agnieszka on his team and to jointly prepare an aggressive push in the development of their project. The circumstances of the conversation are fairly complex but for our discussion we again focus on a small set of issues: the team's chance of success if the team stays together, the team's chance of success if the team splits up, and Dimitri's and Agnieszka's history of past actions. This history is in the background but influences the conversation, and more elaboration on it will follow in later chapters. We can also restrict attention to two types of statements that Dimitri may make in the elevator: statements about the possible scenario where Agnieszka remains on the team – he indeed chooses to make such a statement – versus statements about the possible scenario where the team splits up. In either case, his statement is bound to be fairly bad-tempered, owing to his disappointment about the lost competition. Agnieszka's reaction, in turn, can be simplified to either staying on the team, or not. (Notice also that Dimitri does not have to say anything during the elevator ride. We observe that this may be the wiser course of action, but the subsequent discussions will take it as given that Dimitri is the talker.) The relevant types are Dimitri's and Agnieszka's preferences vis-à-vis these possible statements and vis-à-vis the future actions that may arise in each of the possible scenarios.

The book makes additional restrictions, to narrow things down further.

The first restriction is to consider only single steps of a conversation: there are two people in the conversation – the two interlocutors – where one person says something and the other person reacts.

One may try to apply the book’s analysis to longer conversations with more than one step. To do so, one could simply view everything that is said in the conversation as a single statement. The analogy is vague, however, and it is therefore best to focus on single-step conversations.

Future actions of the interlocutors may, of course, have a foreshadowing effect on the present conversation. [For example, the possibility of telling other children about the encounter with Ralph is on Steve’s mind.](#) The book does not analyze how the interlocutors choose such future action – they are not part of the present situation. It merely considers that future actions are plausibly influenced by what is said now. For a streamlined notation, all future actions are subsumed as parts of the interlocutors’ action sets, even though the interlocutors cannot pick them in the present conversation.

As a second restriction, the book assumes that each person knows their own type. Introspection is assumed as faultless. In contrast, the other person is somewhat mysterious. In formal notation, let  $I_{\theta^j}^i$  be person  $i$ ’s information about person  $j$ ’s type, and let  $\Theta^j$  be the set of possible types for  $j$ . Then, the above assumption is that  $I_{\theta^i}^i = \theta^i$  and  $I_{\theta^j}^i \subset \Theta^j$ .

The third restriction is that the uncertainty about other people’s types is assumed to be statistically independent of the uncertainty about the state of the world. This means that one cannot learn anything about the state from learning about the type, and vice versa. It is another simplification; in its absence the exposition would be cumbersome, with little insight added.

Translating this assumption into a formality is not possible yet, as we have not introduced any formalities about probabilities and beliefs. Once they are introduced, further down in this chapter and the subsequent ones, making the assumption of statistical independence is straightforward. The book’s analysis also assumes that the interlocutors agree to this simplification.

More simplification comes from the fourth restriction: person  $i$ ’s realized utility from the conversation depends *only* on the realized values of the uncertain aspects. These are: the actions of the two interlocutors ( $a^i, a^j$ ) (including their statements in the conversation and everything they do after-

wards), the realized value of the state of the world  $\omega$ , and the person's type  $\theta^i$ .

The fifth restriction is perhaps the most controversial: everyone deals with the uncertainty in a subjectively “rational”, expectation-based way. That is, while the people in the conversation cannot know the uncertain aspects when having the conversation, they form subjective expectations about them and react to these expectations.

Formally, the fourth and fifth restrictions amount to assuming that person  $i$ 's utility from the conversation is given by

$$u^i : A^i \times A^j \times \Omega \times \Theta^i \rightarrow \mathbb{R}$$

(which means that  $u^i$  is a function with arguments  $a^i$ ,  $a^j$ ,  $\omega$ , and  $\theta^i$ ) and that the chosen action  $a^i$  maximizes  $i$ 's subjective expected utility, i.e., she acts to maximize the expected value of  $u^i$  given her subjective expectations.

The book's focus is on these subjective expectations. For brevity, call them *beliefs*. The book simply asks, “What are the beliefs that would justify leading the conversation in the way that people lead it?”

The reader may wonder, why ask such an indirect question about underlying beliefs? Why not ask questions about the conversation itself?

The answer is twofold. First, the book can rely on the existing literatures on decision theory and game theory. These are impressive bodies of literature that describe, among other things, the existence and other properties of beliefs that justify actions. Without going into detail, it is safe to say that the book solidly stands on the shoulders of giants.

Second, a key observation lies at the heart of the book: conversations cannot be correct or false in an unambiguous way, but beliefs can be correct or false. Moreover, one can measure whether they are correct or false.

The next chapters therefore ask whether the beliefs about the uncertain aspects are distorted. For formal notation, the book uses the symbol  $P_x^i$  for person  $i$ 's subjective belief about unknown item  $x$ . It is a probability distribution over  $x$ 's possible values and will depend on information that person  $i$  has. For example,  $i$ 's belief about the state of the world after seeing her information set  $I_\omega^i$  is  $P_\omega^i(\cdot | I_\omega^i)$ , a probability distribution over the elements in  $\Omega$ . As another example,  $P_{a^j}^i(\cdot | I_{\theta^j}^i)$  is  $i$ 's belief about  $j$ 's action  $a^j$ , after



learning information  $I_{\theta^j}^i$  about  $j$ 's type, and it is a probability distribution over the elements of  $A^j$ .

This leads to the final simplifying restriction: common knowledge of the conversation's basic ingredients. The sets of possible actions, types, and states of the world, are assumed to be commonly known by the interlocutors and the information structure that governs who may receive what information is also assumed to be commonly known by them. That is, while each interlocutor does not know what the other knows, they do agree on what this knowledge could conceivably be. Moreover, the fact that each interlocutor reacts to their belief, as specified by subjective expected utility, is also taken as commonly known. The interlocutors do not know each other's beliefs, but know the fact that the other interlocutor reacts to *some* beliefs.

Yet, we note well that beliefs are not identical to information: information is taken as given, but beliefs may be off target. To continue using the notation above,  $I_\omega^i$  is given but  $P_\omega^i(\cdot|I_\omega^i)$  is not. Person  $i$  will, of course, lead the conversation in a way that depends on  $P_\omega^i(\cdot|I_\omega^i)$ . We are therefore interested in measuring it.

Taken together, the above simplifications allow a precise description of the scope of possible belief distortions. This is the book's main subject, whose discussion begins after the next short section.

A quick final comment on notation: to avoid all-too-frequent use of "person  $i$ " and "person  $j$ ", the text henceforth also refers to "us" and "them", where possible.

## Talking and listening

When we talk, the conversation moves *their* beliefs. We can thus influence their actions in a way that improves our utility. When we listen, the conversation moves *our* beliefs: about the issue and about their (the other person's) type. Knowledge about these aspects also informs us about their subsequent actions. Upon listening, we can thus use our improved knowledge to choose a better action ourselves.

Steve's question steers Ralph's attention away from the events underlying his despair (the most relevant issue) to the everyday topic of the ball game.

He also keeps Ralph from thinking about the consequences that arise from the fact that Steve saw him in his present situation. It is noteworthy that Steve does not give Ralph any information but he nevertheless affects his beliefs. The strategy avoids the dangerous terrain, and Ralph sees no reason to change the topic, either. Steve's statement was successful.

Rachel does not have much success with her statement and we may speculate that one reason for it is that the statement is rather impersonal. She does not, in particular, give the governor any motive to follow her funding suggestion. His belief about the political consequences arising from questions about the provenance of art did not change anywhere near as much as would be necessary to choose the action that Rachel hopes to induce.

Not to ignore: there is also a direct utility effect that arises from both our talking and our listening. Talking has value over and above the exchange of information, and so does listening.

Dimitri may feel better after saying what he says, but his statement fails to have the desired effect on Agnieszka's belief: that she may benefit from staying on Dimitri's team. If anything, it has the opposite effect.

Prior to saying something, the direct utility effect from doing so is evident to us: we know what we can say and how much we like saying it (the book's analysis assumes it, anyway – recall that these are simplifications). We also know our type. All other described effects are unknown. We cannot know what they would say if we let them talk. We cannot know how they would react to any statement that we may make. We cannot know the state of the world.

On the basis of our beliefs about these unknowns, we decide how we talk. We also understand that their situation is similar: we know that they, too, form beliefs and react to them. We even know that they know that we form beliefs and react to them. This is indeed a very rational way of conversing.

But rational does not mean optimal. It only means that a conversation follows systematic patterns, which makes it worthwhile to study it. Optimality requires far more: that beliefs are accurate. This is what the 18 questions are all about.