2. Some preliminary remarks

Before examining three popular ideals about how science operates and what constitutes good science and how they fail in the context of climate science, I will first introduce some terms and concepts that will be of relevance throughout this discussion: the epistemic challenges of complex systems, the distinction of context of discovery and context of justification in philosophy of science, and the concept of scientific objectivity. These are not ideals in their own right but rather form 'recurring themes' and are presented here separately from the aforementioned ideals because they are essential to the following discussion in two ways: on the one hand, they either play a constitutive role in the development of these ideals or show how the ideals fall short in the context of modern sciences. On the other hand, a closer investigation of how these motifs 'behave' specifically in the context of climate science will give an indication how the gap left by these ideals can be circumvented. At the end of the next chapter (Chapter 3.4) it will be shown that one element these 'recurring themes' have in common when viewed through the lens of climate science is that they highlight the relevance of the experience or skill that scientists develop through their work. Following from this, a concept of expertise, based on this kind of experience, is discussed in Chapter 4. There it is argued that this concept can – in some public debates about the trustworthiness and reliability of specific scientific research – function as a substitute for the ideals.

2.1 Epistemic challenges of highly complex systems

It seems to be a natural tendency of science to investigate increasingly more complex systems for two interconnected reasons. On the one hand, scientists turn their attention to ever more complicated questions. One way to do so is to examine continuously more complex systems. These, on the other hand, become also more and more accessible to scientific research as the technology advances in a way that creates new instruments to explore these complex systems. Specifically, computer and computer simulations have been significant in tackling complexity in science (Lenhard, 2019).

Although the term is not sharply defined, complexity in science is often loosely understood "as a consequence of numerous independent interacting parts" (Strevens, 2016, p. 696; Weaver, 1948). A complex system usually refers to non-reducible systems with certain characteristics such as non-linearity, emergence, interactivity and path-dependency. The climate system is a complex system par excellence and fulfils, as the next few chapters will show, all these requirements. To assess the epistemic challenges of climate science, the complexity of the system in question is essential. However, I should note here that in the next few chapters I will not continue to assess to what extent exactly the climate system is complex or what particularly defines such a system. Instead when it comes to understanding why climate science cannot fulfil specific expectations about how science is supposed to function, a coarser definition of complexity will be sufficient. The relevant question here is not so much what exactly defines complexity but what follows epistemically from the fact that the climate system – and the computer simulations used to explore it – shows a broad variety of features of complex systems. More specifically, what is particularly significant here is that the complexity has consequences for the question to what extent understanding (of the climate systems and the models) can be achieved. What does that mean? Let's take a look at the particular epistemic challenges of climate modelling.

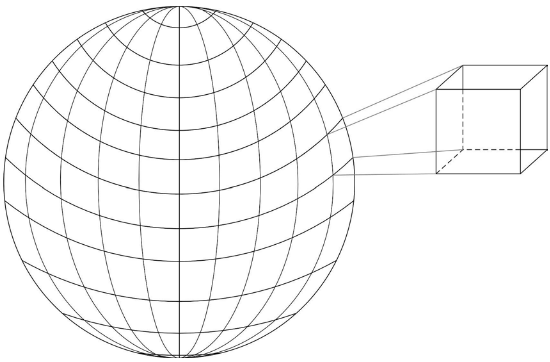
At the core of modern global climate models, so called Earth System Models (ESM), are a number of basic partial differential equations based on well-established principles and laws of physics, such as Newton's second law, thermodynamics and Navier-Stokes equations.[1] Historically, these mathematical descriptions of the dynamics of the climate system arise out of what are today referred to as the *primitive equations*, first developed by Vilhelm Bjerknes in 1904.[2] However, because these equations cannot be solved analytically, climate

1 For a more elaborated history of the development of climate models see Weart (2010); Gramelsberger (2011) and especially Edwards (2010).

2 To be more precise, Bjerknes proposed that the dynamic of the weather system could be described by these seven equations:
"1. The three hydrodynamic equations of motion. These are differential relations among the three velocity components, density and air pressure.

scientists have resorted to computer simulations. This is done by discretising these analytically unsolvable equations to calculate the state of the system step by step in specific intervals of time. In the case of global climate models, this means that the globe gets covered by a virtual three-dimensional grid (see Figure 1).[3] Information about climate variables such as temperature, pressure, humidity and wind are then determined for every grid cell at discrete time steps and is shared with neighbouring cells. Although this is often referred to as a *numerical solution*, the transformation of the analytical equations is only an approximation and the source of some uncertainties in climate modelling (see Chapter 3.3.3).

Figure 1: Discretisation grid for a climate model



---

2. The continuity equation, which expresses the principle of the conservation of mass during motion. This equation is also a differential relation, namely between the velocity components and the density.
3. The equation of state for the atmosphere, which is a finite relation among density, air pressure, temperature, and humidity of a given air mass.
4. The two fundamental laws of thermodynamics, which allow us to write two differential relations that specify how energy and entropy of any air mass change in a change of state" (Bjerknes, [1904] 2009, p. 664).

3 Typical depictions of climate models show cube-shaped grid cells (like in Figure 1). Those have the disadvantage that they differ in size in a spherical system (i.e., the closer to the pole, the smaller they get). To tackle this inconsistency in grid size many contemporary climate models now have other grid shapes, e.g., icosahedral grids (Zängl et al., 2015).

Processes deemed too complex and/or (more importantly) taking place on a scale that is below the grid size are integrated into the models in form of parametrisation (McFarlane, 2011).[4] Processes, such as cloud formation, radiation, atmospheric convection but also tree growth, are happening on a scale that is too small and/or are not well enough understood to express in terms of physical laws to be represented in a resolved way at the grid-scale level. Nonetheless, these processes are considered to be too important to the climate to be just left out of the model. So to include these kinds of processes, scientists developed parametrisations that serve as substitutes for sub-grid level processes in the form of functions of above-grid-level variables.

Parametrisations are sometimes called being "semi-empirical" (for example Edwards, 1999, p. 449; Parker, 2018) as well as having been noted to include "non-physical" elements (for example Winsberg, 2018, p. 48). On the one hand, parametrisations are usually only partially based on insight into and physical description of the climate system but also on approximations gained from observations. On the other hand, usually some parameters in parametrisations are not physically constrained and do not have a direct equivalent in the actual target system. Thus, they are not set to represent a real-world 'true' value but a 'best' value. This makes parametrisations, both in the necessity for them and the way they are constructed, "artifacts of the computation scheme" (Winsberg, 2018, p. 49). Creating parametrisation schemes can be a highly complex process and there are frequently different options of how to parametrise the same process. This makes parametrisation a critical source for uncertainties in climate modelling (see Chapter 3.3.3).

The development of global climate models is an intricate and time-consuming task. A typical climate model of the scale of an *Earth System Model* (ESM), which are currently considered the "state-of-the-art" climate models, consists of hundreds of thousands lines of code, written and further developed by several generations of scientists – often over more than a decade. Because of this, these kinds of models are usually not built from scratch but incorporate bits and pieces from previous generations of models (Knutti et al., 2013). It is also not uncommon for scientists to 'borrow' parts of or even whole model components from other institutions and modify it so that it fits with their model

4 Climate scientists commonly differentiate between the physics and the dynamics of an atmospheric model. The former refers to those processes that have to be parametrised and the latter to the dynamical processes that can be described in a resolved form.

when they lacks that specific expertise (Alexander and Easterbrook, 2015, pp. 1227–1228).

Structurally, a climate model is constructed out of several different components. A typical ESM consists of, for instance, an atmosphere, an ocean, a sea ice, a land and vegetation component.[5] The exchange of relevant variables at the border regions of the different model components is facilitated via a so-called *coupler*. A coupler compensates for the difference in resolution of different model parts and supports the exchange of information between the various model components.

Climate models are "highly modular" (Winsberg, 2018, p. 50). There are a number of different ways the coupling of the different model parts can be constructed. For example, a study by Alexander and Easterbrook (2015) shows that models coming from American modelling centres often have a "star-shaped" structure, where every component is connected directly through a coupler to the other components. Models from European climate modelling institutes tend more towards a "two-side" structure, where only the atmosphere and ocean component are directly connected to the coupler. The land, sea ice and other possible elements are integrated into either the ocean and atmosphere component requiring that the resolution of minor components is the same as the atmosphere or ocean model, respectively (Alexander and Easterbrook, 2015, p. 1225). In a similar fashion there are usually a variety of ways in which specific climate processes are integrated into a model. Depending on the resolution of the model, particular processes need to be parametrised or can be integrated into the model in a resolved way. Further, because of limited computing power, time-constraints and previous modelling decisions not all processes can be represented equally well in a model. How a model component

5 Alexander and Easterbrook note that there are certain analogies between how the expertise in climate modelling facilities, the computer model and nature itself is structured: "The boundaries between components in an Earth system model represent both natural boundaries in the physical world (e.g., the ocean surface) and divisions between communities of expertise (e.g., ocean science vs. atmospheric physics). The model architecture must facilitate simulation of physical processes that cross these boundaries (e.g., heat transport) as well as support collaboration between knowledge communities within the work practices of model development (e.g., to study climate feedbacks)" (Alexander and Easterbrook, 2015, p. 1224). However, it should be mentioned that there are, of course, also limitations to these similarities and one should not assume that nature 'functions' in the same way as the models (specifically with respect to parametrisations).

is structured, whether a process is integrated directly or through a parametrisation, which processes and variables get preferential treatment leaves many different options of how to build a climate model (Parker, 2018). In practice, how these decisions are made often depends on different modelling traditions and trade-off considerations (see Chapter 3.1.3). That is, that there is no such thing as a 'universally agreed upon construction manual' for climate models.

One might now think that, because ESMs are composed of many different parts, this means that, even though the models as a whole are complex and large structures, the quality of the models can still be assessed relatively easily by testing the different model parts separately. However, there is a high interdependency and exchange between various elements of the models. In practice, the components of the model are, of course, also tested extensively and calibrated on their own but will inevitably perform differently when put together (Hourdin et al., 2017, p. 591). Lenhard and Winsberg have called this the "fuzzy modularity" of climate models:

> In sum, climate models are made up of a variety of modules and sub-models. [...] And it is the interaction of these components that brings about the overall observable dynamics in simulation runs. The results of these modules are not first gathered independently and then only after that synthesized. Rather, data are continuously exchanged between all modules during the runtime of the simulation. The overall dynamics of one global climate model is the complex result of the interaction of the modules—not the interaction of the results of the modules. For this reason, we like to modify the word "modularity" with the warning flag 'fuzzy': due to interactivity, modularity does not break down a complex system into separately manageable pieces. (Lenhard and Winsberg, 2010, p. 256)

The high interdependency and continual exchange between the different parts of the model can also give rise to compensating effects to the extent that some feature of one model part will interfere with another element of the model in such a way that it makes up for some particular shortcomings of the individual model component.

Further, climate modelling requires tuning. Tuning is a process at the end of a model- or submodel-construction cycle, where a few parameters (for example, parameters concerning cloud or surface albedo properties) are adjusted so that the model behaves in a way scientists consider to be realistic (Mauritsen et al., 2012). Although there is some common consensus about some general aspects of the tuning process and goals, how a specific model is tuned depends considerably on the objectives and preferences of the modelling group in question (Chen et al., 2021, pp. 217–218).[6]

As will be discussed further in Chapter 3.1.3, tuning is always a question of trade-off, meaning that one cannot tune a 'perfect' model in respect to every variable. Climate scientists have also voiced concern that tuning models for 20th century warming may inadvertently lead to and mask compensating errors (e.g., Mauritsen et al., 2012): that is, that the improved model performance is rooted in hidden structural problems.

Lenhard and Winsberg have famously argued that all this makes climate models "analytical impenetrable in the sense that we have been unable, and are likely to continue to be unable, to attribute the various sources of their successes and failures to their internal modeling assumptions" (2010, p. 261). With reference to William Wimsatt's concept of "generative entrenchment" (2007, p. 133), they argue that the models' "layered history" (Parker, 2018) intricately influences the performance of the model. On the one hand, many aspects of climate-model development are not fully epistemically constrained. That means that there is usually more than one option how to integrate particular features of the climate system into a model. On the other hand, what kind of steps can or will be taken next in the continual development of a model is constricted by previously made modelling decisions. This *path-dependency* of the models means that climate modellers are not infrequently limited in their modelling choices by deliberations of previous generations of scientists and they will, in turn, further constrain the options and strategies available in the future. Further, there are always trade-offs to be made and, because there is a limit to computing power and time available, not all relevant climate processes can be represented equally well within the model (see Chapter 3.1.3).

This goes hand in hand with what Lenhard and Winsberg, with reference to Andy Clark (1987), call *kludging*. Kludging describes the piecemeal construction of complex computer programs that is effective but, nevertheless, "botched together" (Clark, 1987, p. 278) to the extent that as it does not follow a coherent construction plan. This can contribute to a lack of or reduction in analytical

6 A common target of tuning, for example, is that the mean equilibrium temperature the models display is in agreement with observations. A frequently used strategy to achieve this is to regulate the top-of-the-atmosphere energy balance through adjusting, e.g., cloud parameter. The adjustment of these parameters is important as global warming essentially is an energy imbalance (i.e., a mismatch between the incoming and outgoing energy) of the system (Hourdin et al., 2017).

understanding of the program. Considering what we have learnt so far about the specific characteristics of climate model developing – multiple generations continually adjusting the project, often driven by pragmatic considerations – it is easy to see why kludging would both occur here and further affect the access to analytical understanding.

There is some disagreement among philosophers to what extent these limitations in analytical understanding are 'here to stay' or can be overcome (at some point). Lenhard and Winsberg (2010) argue that climate modelling is affected by a strong form of *confirmation holism* often preventing a precise attribution of the source of error in a model. Parker (2018), however, notes that there has been some progress in gaining analytical understanding in recent years (e.g., through finding so-called *emergent constraints* (see Chapter 3.3.3.4)). It is, nevertheless, a difficult and laborious undertaking.

This gives us a first glimpse of the epistemic challenges of the high complexity of the climate system and the models climate scientists employ to tackle this complexity. It also indicates why the failure of certain ideals of science becomes so apparent in the wake of the difficulties described above in gaining analytical understanding. All these aspects will be discussed further in the following chapters.

It should be pointed out here that all sorts of different types of models are used in climate science. Although I will primarily focus on ESMs (*Earth System Models*), which are the state-of-the-art global climate models, and their predecessors AOGCM (*Atmosphere-Ocean General Circulation Models*),[7] climate scientists also make use of regional models to estimate climate development on a local level. Global models of various complexities lower than that of an ESM are also frequently applied in those cases where a reduced demand in computing power is advantageous (Chen et al., 2021, pp. 218–219). Many of the issues and philosophical challenges to be discussed in the next chapter also apply to these models. To understand the impact of climate change and options of mitigation, scientists also employ other types of models such as *Integrated Assessment Models* (IAM), which will also not be discussed here.

7 Compared to AOGCMs, ESMs also include biochemical processes (Chen et al., 2021, p. 181).

2.2 Discovery and justification: the DJ distinction

When Hans Reichenbach published his landmark book *Experience and Prediction* in 1938, his primary intention might have been to introduce himself and his brand of philosophy to the American philosophy of science community after he had to emigrate from Germany (Howard, 2006, p. 7). But the first chapter also spelt out a concept that would impact the (self-)perception of philosophy of science for the rest of the century: the distinction between the context of discovery and the context of justification (usually abbreviated to *DJ distinction*). The DJ distinction significantly reduces the scope of philosophy of science and, thereby, has had subsequently a (sometimes somewhat concealed, sometimes more obvious) influence on many discussions and controversies in philosophy of science. It should, therefore, not be surprising that we will see this distinction popping up in several places throughout the discussion of specific popular ideals about how science ought to operate in Chapter 3. The reason why I introduce this concept and some modern interpretations here separately from any distinct ideal is twofold: first of all, the DJ distinction is constitutive not just for one but two of these ideals. Secondly, as we will also see in the next few chapters, in the context of complex computer simulation, as they are used in climate science, even a "lean version" (Hoyningen-Huene, 2006) of the DJ distinction cannot be upheld as it is difficult to fully separate the evaluation of these simulations from their history (see Chapter 3.4.2).

But before we turn our attention to any modern interpretation of the DJ-distinction, what was Reichenbach's original reasoning? His overall claim is that there are three tasks epistemology has to tackle: the *descriptive task*, the *critical task* and the *advisory task*.

The first task of epistemology is "giving a description of knowledge as it really is" (Reichenbach, 1938, p. 3). This, however, does not mean that the epistemologist should be concerned with describing all and any of the thoughts the scientists actually had before coming to a particular conclusion. That sort of reconstruction of a scientific thought process falls, argues Reichenbach, into the realm of psychology. Instead, the job of the epistemologist is to perform a "rational reconstruction" (Reichenbach, 1938, p. 5):

> What epistemology intends is to construct thinking processes in a way in which they ought to occur if they are to be ranged in a consistent system; or to construct justifiable sets of operations which can be intercalated between the starting-point and the issue of thought processes, replacing the real

> intermediate links. Epistemology thus considers a logical substitute rather than real processes. (Reichenbach, 1938, p. 5)

Hence, what Reichenbach has in mind is a logical reconstruction of an ideal thought process somewhat comparable to the reconstructed thought processes which scientists themselves publish in scientific journals to communicate their findings to their peers.[8]

The second task of epistemology is the *critical task*. It overlaps in some ways with the *descriptive task* but must, as Reichenbach insists, be viewed separately, for its central objective is not just to describe but to criticise "the system of knowledge […] in respect of its validity and its reliability" (Reichenbach, 1938, p. 7). Besides examining the logical basis of science, another main function of the critical task is to point out "volitional decisions" (Reichenbach, 1938, p. 9). Reichenbach acknowledges that the scientific process includes many instances in which the next step cannot be determined by logical deliberations alone. Instead, scientists routinely have to make methodological decisions between two or more equally good options. Detecting and disclosing these "volitional decisions" is "one of the most important tasks of epistemology" (1938, p. 9). This includes the specification of conventions (for example, measuring units) and "volitional bifurcations" (Reichenbach, 1938, p. 10), compared to conventions these are decisions which do not result in equivalent systems.

The third task is the *advisory task*. Unlike what one might surmise from the name, the advisory role of epistemology must, according to Reichenbach, be curtailed to the bare minimum. That means the epistemologist must refrain from giving direct advice which decisions to take. Instead, Reichenbach argues that the only appropriate part for philosophers in the decision-making process in science is to point out different available options:

> We may therefore reduce the advisory task of epistemology to its critical task by using the following systematic procedure: we renounce making a proposal but instead construe a list of all possible decisions, each one accompanied by its entailed decisions. So we leave the choice to our reader af-

8 Reichenbach, however, stresses that common scientific writing is not precise enough for the kind of logical reconstruction epistemologists should aspire to: "For scientific language, being destined like the language of daily life for practical purposes, contains so many abbreviations and silently tolerated inexactitudes that a logician will never be fully content with the form of scientific publications" (Reichenbach, 1938, p. 7).

> ter showing him all factual connections to which he is bound. (Reichenbach, 1938, p. 14)

Reichenbach emphasises that the role of philosophers here is at best to make a "proposal" by calling attention to the advantages and disadvantages of a decision, not a "determination of truth-character" (Reichenbach, 1938, p. 13). One particular concern for philosophers in this context, Reichenbach points out, are what he calls "entailed decisions" (Reichenbach, 1938, pp. 13–16), that is, tracking and identifying the consequences of decisions. In this situation, the philosopher is, according to Reichenbach, in the position to show how specific disputed decisions logically follow from already well-established ones.

Reichenbach, thus, here sets clear boundaries for the scope of philosophy of science. Its work should be restricted to the reconstruction and evaluation of scientific arguments from a logical point of view. The advisory role of epistemology is limited to pointing out different options and does not extend to an interference in the actual decision-making processes.

Setting the scope of philosophy of science in terms of a distinction between the realm of discovery and the realm of justification quickly became prevalent in philosophy of science for some time. Beginning in the 1960s and at the start of the decline of the dominance of logical empiricism in philosophy of science, some philosophers began to voice criticism. They argued against the omission of the dimension of discovery and history from philosophy of science (Schickore and Steinle, 2006a). Instead, they argued for a "logic of discovery" (Nickles, 1980), for including historical analysis and for returning scientific practice and experimentation back into the philosophical limelight (see Chapter 3.2.1).

Paul Hoyningen-Huene (2006) comes to the conclusion that some of the resistance against the DJ distinction can be traced back to some confusion that arose from the fact that by the mid-century there were several different versions of the DJ distinction at play often muddled together. According to Hoyningen-Huene, this led to a situation where in respect to the DJ distinction it was "not clear what exactly is stated by its defenders and what exactly is attacked by its critics. Eventually, all parties, growing frustrated, turned away from the discussion" (Hoyningen-Huene, 2006, p. 119). As the DJ distinction will be a recurring topic in the next chapter, it is worthwhile to take a closer look at some of the different versions of the DJ distinction that Hoyningen-Huene identifies (2006, pp. 120–123). Historically, he claims there are the following five different variations of the DJ distinction in the literature of the 1960/70s:

1. two temporally distinct processes
2. the historical discovery process versus the specific justification methods
3. an empirical versus logical process
4. a disciplinary distinction
5. a differentiation in respect to different questions asked

First, there is the temporal distinction. Here discovery and justification are seen as two processes taking place one after the other. Initially, there is a discovery (here the term can be stretched to include inventions), which is followed by a justification process. This definition of the DJ distinction does not hold up to actual scientific practice. The second definition focuses on a distinction between discovery processes and justification methods. That is, there is a "contrast between the factual historical process and methods, considerations, procedures, etc. that are relevant to justify or to test knowledge claims" (Hoyningen-Huene, 2006, p. 121). Hoyningen-Huene argues that this can be either interpreted historically, running into the same problem as the first definition or normatively; meaning that "historical processes (of discovery) are described, whereas claims of justification or testing are normatively evaluated" (2006, p. 122). One can specify this version further by defining discovery as an empirical process and justification as a logical process. Following this distinction, one might also attribute different disciplines to the two categories. History, science and psychology of science are considered to be methodologically empirical, whereas philosophy of science is methodologically logical. The fifth version of the DJ distinction that Hoyningen-Huene identifies in the literature is that of two different and distinct questions being asked "such as 'What has happened historically during this discovery?' versus 'Can a statement be justified? Is it testable?'" (2006, p. 123). Hoyningen-Huene argues that specifically versions 1–4 (historically) are often merged into one,[9] leading to a logical em-

9 The implication of this, Hoyningen-Huene asserts, is that "*a rational disagreement about justification is conceptually impossible*" (2006, p. 124). It is assumed that the sphere of discovery can only be subjected to empirical investigations. Thus, there is no place for philosophy and there is no such thing as a "logic of discovery". The mixing of different versions of the DJ distinction also leads to the assumption that the only form of justification is logical and it is the 'business' of philosophy of science to evaluate it, Hoyningen-Huene argues. That means that any conflict on grounds of questions of justification

piricists' understanding of what constitutes philosophy of science (2006, pp. 123–124).

Hoyningen-Huene himself, subsequently, argues for what he calls a "lean" version of the DJ distinction that includes elements of version 2 and 5 (2006, pp. 128–130). The general idea is that there is a difference between a *descriptive* and a *normative perspective*:

> From the descriptive perspective, I am interested in facts that have happened, and their description. Among these facts may be, among other things, epistemic claims that were put forward in the history of science, that I may wish to describe. From the normative or evaluative perspective, I am interested in an evaluation of particular claims. In our case, epistemic claims, for instance for truth, or reproducibility, or intersubjective acceptability, or plausibility, and the like are pertinent. Epistemic norms (in contrast to, say, moral or aesthetic norms) govern this evaluation. By using epistemic norms we can evaluate particular epistemic claims according to their being justified or not. (Hoyningen-Huene, 2006, p. 128)

Hoyningen-Huene argues that what makes this version "lean" is that it merely distinguishes between two different perspectives. Thereby, one does not have to make any additional assumptions, such as that there cannot be any overlap between the two spheres both in a categorical and a procedural sense (see also Chapter 3.4.2).

The context distinction, as envisioned by Reichenbach and, subsequently, the logical empiricists, limiting specifically the scope of philosophy of science to a logic of justification is now seen as out-dated by the vast majority of philosophers of science. But the aftermath is still felt, argue Schickore and Steinle:

> In recent years, philosophers have rarely directly addressed, let alone attacked the distinction. But this does not mean that the distinction has been rendered irrelevant or that it has been successfully refuted. On the contrary, the legacy of earlier advocates of the distinction is still effective, and the distinction continues to delineate the scope [*sic*] philosophy of science. (Schickore and Steinle, 2006a, p. ix)

---

can only emerge with respect to either an error or differences in conventions. However, as differences in conventions are not considered to be "epistemically substantial disagreements", the analysis of potential errors in the logical justification "is a one-person-game" (Hoyningen-Huene, 2006, p. 124).

Schickore and Steinle see both the separation and lack of exchange between philosophy of science, history of science and science studies as well as the lack of interest from the analytically orientated philosophy of science, only rarely integrating knowledge from the other two disciplines, as a lasting sign of this (Schickore and Steinle, 2006a, pp. ix–x). And, as already mentioned, it also had a prolonged influence on how philosophers, scientists[10] and the public have idealised or still idealise the inner workings of science. For this reason, the DJ distinction will also be crucial for understanding the history of two of the three ideals of how science does and should operate discussed in the next chapter. On the one hand, the notion that science should disregard sociological and psychological aspects of science was a constitutive element to the ideal of value-free science (see Chapter 3.1). On the other hand, the DJ distinction was also a significant contributing factor to the neglect of the experimental part of science by philosophy of science during much of the 20th century, as the experiment is traditionally seen as mostly an element of the discovery side of science that Reichenbach has declared to be of no philosophical interest (see Chapter 3.2).

In the context of climate modelling, it will also become apparent that the DJ distinction is problematic even in the weaker form that Hoyningen-Huene proposes. As will be discussed at the end of the next chapter (Chapter 3.4.2), in the context of climate science, it is no longer possible to fully separate the evaluation and justification of models as well as the techniques employed in their construction and evaluation from their own history. This highlights the relevance of the experience climate scientist develop in working with the models, which will explain why a conception of expertise rooted in this experience can be a successful way for those who are outside of the scientific process to assess when to be sceptical about claims made by apparent scientific 'experts'.

10 Interestingly, Schickore and Steinle argue that the DJ distinction is still very much alive in science itself: "Remarkably, today the distinction is most explicitly discussed in the sciences themselves. In methodological introductions of science textbooks, it shapes the regulations for scientific research. These textbooks employ a particular version of the distinction, namely the context distinction temporally understood in combination with the hypothetico-deductive (H-D) model of scientific research" (Schickore and Steinle, 2006a, p. ix). Schickore and Steinle however also note that, when scientists perceive this to be the actual view that philosophers of science have about science, it can become a point of conflict and lead scientists to criticise philosophers for having an oversimplified concept of science.

2.3 A few words about objectivity

Since *scientific objectivity* was first established within science in the middle of the 19^{th} century as a goal that scientists should strive towards (Daston and Galison, 2007, p. 27), it has become a rarely questioned concept in science. These days, the term *objectivity* is almost used like a 'magic word' in science. It is a word invoked by scientists, science communicators but also philosophers whenever they want to stress that science is something 'special', something that sets science apart from pure opinions. The *objectivity* of either research, researchers or lack thereof is a claim that is quickly resorted to in public debates about science. When something, a research result or a person or a method, is declared as *objective* in science, it is supposed to be a sort of seal of approval as Reiss and Sprenger note:

> Using the term "objective" to describe something often carries a special rhetorical force with it. The admiration of science among the general public and the authority science enjoys in public life stems to a large extent from the view that science is objective or at least more objective than other modes of inquiry. (Reiss and Sprenger, 2017)

In the context of public debates about science, the apparent *objectivity* of science is alternately used to 'prop up' scientific research results as irrefutable and significant or to undermine the work of the scientists by proclaiming that they are not 'objective'. At the same time, scientists themselves frequently 'conjure up' the term when they describe their own work or methods.

While the term *objectivity* is used often as the 'ultimate' signifier of 'good science', an in-depth examination of the use of the word *objectivity* in science quickly reveals that this is far from as clear-cut as the confidence, with which the term is applied. In fact, as Heather Douglas argues, *scientific objectivity* is "among the most used yet ill-defined terms in the philosophy of science and epistemology" (Douglas, 2004, p. 453).[11] In fact, a quick look at the literature also shows that there is not even a clear consensus among philosophers of science about the exact amount of definitions.

---

11 This also applies to the word *objectivity* in its historical evolution. Even though objectivity as an objective of science only has existed since approx. 1860, the history of the term *objectivity* itself goes further back and has quite a "somersault history" (Daston and Galison, 2007, p. 27), having changed its meaning almost to the contrary since first appearing in European languages.

Reiss and Sprenger, for instance, state that, in principle, there are two distinct basic concepts of scientific objectivity: *product objectivity* and *process objectivity*:

> According to the first understanding, science is objective in that, or to the extent that, its products — theories, laws, experimental results and observations —constitute accurate representations of the external world. The products of science are not tainted by human desires, goals, capabilities or experience. According to the second understanding, science is objective in that, or to the extent that, the processes and methods that characterize it neither depend on contingent social and ethical values, nor on the individual bias of a scientist. (Reiss and Sprenger, 2017)

The authors emphasise that particularly *process objectivity* comes in a variety of forms depending on what kind of scientific *process* (e.g., the structural organisation of science and methods of measuring) is meant to evoke objectivity.

In a similar, but somewhat more specific, fashion, Martin Carrier (2013, 2010) also distinguishes two different and contrasting concepts of scientific objectivity. He argues that concepts "of scientific objectivity are governed by two ideal types, namely, objectivity as adequacy to the facts and objectivity as reciprocal control" (Carrier, 2010, p. 207). He traces the former back to Francis Bacon, according to whom the ideal scientists is detached and neutral (Bacon, [1620] 1863; Carrier, 2013, p. 2549, 2010, pp. 207–208). The second meaning of objectivity is pluralistic in its approach. From this point of view, objectivity is reached through a diversity of points of view. This understanding of scientific objective has been popularised by Helen Longino (1990) but Carrier also sees familiar elements of this approach in the works of Karl Popper and Irme Lakatos (for a more detailed discussion of the pluralistic approach to objectivity and values, see Chapter 3.1.1).

Contrary to these dualistic definitions of *scientific objectivity*, other philosophers have argued that a more differentiated categorisation is more appropriate. While Megill (1994) argues that there are "four senses of objectivity" in general: *absolute*, *disciplinary*, *dialectical* and *procedural*, Heather Douglas (2004) finds at least eight different use cases of the phrase *scientific objectivity* divided into three different modes: the first type of objectivity is defined by the interaction of the scientists with the world. Secondly, there is an understanding of objectivity that is characterised through the personal (the value-related) reasoning process of the scientist. The third type of objectivity is based on the social structures and procedures of science.

The different and contrasting interpretations and categorisations of the phrase *scientific objectivity* indicate that the term is "not logically reducible to one core meaning" that simply (Douglas, 2004, p. 455).

It, therefore, seems prudent to follow Douglas here, who argues that these different types of objectivity are neither necessarily reducible nor incompatible. On the contrary, there is often more than one meaning at play when we call something *objective*.[12] Furthermore, the way we use the phrase *scientific objectivity* might change in the future. Certain definitions might be found wanting and others might newly emerge. We are by no means "finished developing the term" (Douglas, 2004, p. 468).

This diversity of meanings should also be kept in mind when we come across the term *scientific objectivity* in the next chapter. The simultaneously wide but not very well defined application of the term is reflected in the context of public climate-science debates. It is, therefore, not surprising that the term *objectivity* will also return in the following discussion about traditional ideals about science which are in conflict with how modern science actually functions. The most striking occurrence of this is in the context of the value-free ideal because in the public discourse objectivity is often used synonymous with the value-freeness of science (see Chapter 3.1). In the next chapter it will become clear why science cannot be *objective* in this sense.[13] Further, the phrase *objectivity* will reappear in connection with the ideal that observations in science can provide irrefutable evidence for or against a theory. In this context observations are commonly treated as *objective* in the sense that they are considered to be not open to interpretation (see Chapter 3.2). Contrary to this, many philosophers of science agree that observations have to be treated as theory-laden and theories are underdetermined by data. Particularly in climate science where a large amount of data has to be dealt with, models are usually considered to be "data-laden" and observations "model-filtered" (Edwards, 1999).

12 This, Douglas argues, also applies to the term *subjectivity* which has a considerable and irreducible variety of interpretations. Nor can subjectivity be seen as just the opposite to objectivity: "subjectivity is not just the lack of objectivity, and objectivity is not just the overcoming of subjectivity. Both are rich concepts, elements of which may be placed in stark opposition to each other" (Douglas, 2004, p. 470).

13 For a more successful application of a concept of scientific objectivity (following Longino, 1990, 2002) to aspects of climate science, see Leuschner (2012a) and Chapter 3.1.3.4.

At the same time the word *objectivity* is also used by climate scientists themselves. Here the term usually has a much narrower application and is commonly located in recurring debates about whether so-called *subjective* 'manual' procedures can be substituted by *objective* mathematical procedures. Here, *subjective* has a somewhat negative connotation. There are, for instance, discussions about whether the process of tuning can be made more objective by implementing an automated process of "find[ing] optimal sets of parameters with respect to certain targets" (Mauritsen et al., 2012, p. 16). In a similar vein, certain methods of data-processing (see Chapter 3.2) and procedures of model intercomparison (see Chapter 3.3), applied to study uncertainties in the models, are sometimes described as *objective* because there is an automated, mathematical element to the method. In that sense, the use of the term *subjective* in describing certain features of scientific methods requiring the specific skill and experience that scientists develop through their work will also be important to understanding the rising relevance of this experience to epistemological questions, as will be discussed in Chapter 3.4.