Imbalance and the State of Research

Emergent Challenges to Scientific Independence & Objectivity

Von der Philosophischen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover zur Erlangung des Grades
Doktor der PHILOSOPHIE
Dr. phil.
genehmigte Dissertation von
David Hopf
Erscheinungsjahr 2020

Referent: Prof. Dr. Torsten Wilholt
Korreferentin: Prof. Dr. Anke Büter
Tag der Promotion: 11.06.2020

Acknowledgments

This dissertation owes its existence not just to me, but to a great many people to which I am greatly indebted:

First of all, I want to sincerely thank my first supervisor Dr. Torsten Wilholt, especially for our many and regular meetings both in his office and in his home; not only for his support and advice concerning philosophy and career but also for the coffee. A great many thanks also to my second supervisor Dr. Anke Büter, who offered many insightful comments and criticisms on the dissertation.

I was lucky to have been made associate member of the DFG research training group GRK 2073; without this opportunity, work and life during this project would have been greatly impoverished in every respect. First, let me thank the many senior members of the group, who commented on my project at the colloquia and other events. I want to give special thanks to Dr. Saana Jukola, Dr. Anna Leuschner, and Dr. Dietmar Hübner, all three of whom gave me additional feedback and advice in individual meetings.

But I am also hugely indebted to my fellow PhD students working with me in the GRK offices in Hannover, who—be it during coffee breaks, in talks over lunch or really at any time of the day—have persistently contributed their insights to my work, but also generally made this PhD project a less solitary and thus much more enjoyable affair. From this group, Daria Jadreškić and Julia Pfeiff deserve special mention as fellow comrades who fought with me through the last months of finishing our respective dissertations. I also want to give additional thanks to Dr. Stefano Canali, who, with his frequent advice on both form and content, helped carve this project at its joints. Special thanks also to Enno Fischer, with whom I shared an office—or rather, two of them in sequence—and many an afternoon discussion.

Similarly, I am very grateful to all my other fellow GRK members working at the institute building in Hannover—but also to all the other colleagues I encountered there—and at Bielefeld University, who I regret to have met less frequently, but who offered no less help or camaraderie when I did.

Not only I but all members of the group can count themselves lucky for the constant support in all matters by our coordinator Dr. Rafael Ugarte Chacón, who made sure that all aspects of work at the GRK smoothly ran together.

I also want to give thanks to the people who offered very helpful feedback on the penultimate draft of this dissertation:

First, to my mother, who did not only help with proof-reading but had an open ear to all my worries throughout the last three years. At this point, I also want to thank the rest of my family—my father, my brothers, my niece, my aunt, my uncles, and my dear grandmother—for their patience during a time where I was not always as available as they deserve.

Second, to Dr. Tanja Rechnitzer, who always has been a great friend but also a mentor in all things philosophical since my very first days at university.

Third, to Christoph Schmidt, who offered to read this text of his own accord, and who shall also have to stand in for all the other friends that helped maintain a balance between work and life after hours.

Last but not least, I am also very grateful to the Volkswagen Foundation, which provided the funding for the collaborative project "The Independence of Research as a Multilevel Problem: Interdisciplinary and Methodological Challenges" in which I was employed and thus enabled work on this dissertation in the first place. My thanks also go out to the other researchers involved in that research group, who, with their very interesting empirical projects, heavily influenced my initial ideas and continuously provided an important perspective beyond philosophical thought.

Zusammenfassung

In dieser Dissertation entwickle ich einen konzeptionellen Rahmen für die Analyse von Problemen mit der Unabhängigkeit und Objektivität des Forschungsstandes. Ich vertrete die These, dass wir einen neuen Ansatz brauchen, der über die bestehenden Vorstellungen von wissenschaftlicher Objektivität hinausgeht. Denn wenn wir uns mit dem Stand der Forschung befassen, müssen wir nicht nur die Qualität einzelner Ergebnisse berücksichtigen, sondern auch das Problem der Unausgewogenheit, d.h. die Vernachlässigung oder das Überangebot bestimmter Arten von Forschung im Vergleich zu anderen. Im ersten der drei Teile der Dissertation definiere ich den Begriff ”Forschungsstand” und stelle die Idee von Unausgewogenheiten im Forschungsstand vor. Letzteres Konzept geht auf die Diskussion verschiedener Beispiele aus der Literatur zurück. Im zweiten Teil analysiere ich drei zentrale Begriffe, die helfen können, zu erklären, warum diese Fälle als Probleme für die Unabhängigkeit und Objektivität der Wissenschaft angesehen werden können und sollten: epistemische Vertrauenswürdigkeit, Produktivität und Gerechtigkeit. Diese drei normativen Kriterien und ihre Wechselbeziehungen bilden die allgemeine Struktur des konzeptuellen Rahmens. Im dritten Teil operationalisiere ich die Kriterien, um zu zeigen, dass und wie sie auf die im ersten Teil diskutierten Fälle angewendet werden können und diskutiere die Ergebnisse. Ich schließe mit einer Erörterung der Implikationen für Konzepte höherer Ordnung wie Bias, der Unabhängigkeit der Wissenschaft und Objektivität. Ich lege dar, dass wir, wenn wir die Anwendung dieser Konzepte auf den Stand der Forschung betrachten, neben rein epistemischen Erwägungen insbesondere auch die soziale Verantwortung der Wissenschaft berücksichtigen müssen.

Schlagworte: Wissenschaftliche Objektivität; epistemisches Vertrauen; soziale Verantwortung der Wissenschaft

Abstract

In this dissertation, I provide a conceptual framework for the analysis of impediments to the independence and objectivity of the overall state of research. I argue that we need a new approach beyond existing conceptions of scientific objectivity. This is because—when concerned with states of research—not only do we have to account for problems with individual findings but also with imbalance, that is the neglect or overabundance of specific types of research relative to others. In the first of the three parts of the dissertation, I define the concept "state of research", and introduce the idea of its imbalance. The latter concept is based on a discussion of various examples from the literature. In the second part, I analyze three major concepts that can help explain why these cases can and should be considered problems for the independence and objectivity of science: epistemic trustworthiness, productiveness, and justice. These three normative criteria and their interrelations form the general structure of the conceptual framework. In the third part, I operationalize the criteria to show that and how they can be applied to the cases introduced in the first part, and present the results. I conclude with a discussion of the implications for higher-order concepts such as bias, the independence of science, and objectivity. I argue that when we look at the application of these concepts to the state of research, in addition to purely epistemic considerations, we also have to emphasize the social responsibility of science.

Keywords: scientific objectivity; epistemic trust; social responsibility of science

Contents

1 Introduction **13**

1.1 Overview 15

1.2 Research Questions & Approach 18

Part I Imbalance in the State of Research (SoR) **23**

2 What is the SoR? **25**

2.1 The SoR and its Reconstruction 27

2.2 Aspects of SoRs 36

2.3 Relevance 43

3 What is Imbalance? **59**

3.1 Imbalance of the SoR in the Literature 62

3.1.1 Lacunae in the SoR 62

3.1.2 Overabundance in the SoR 68

3.1.3 Categories of Imbalance 72

3.2 The Concept of Imbalance 75

3.2.1 A Balanced SoR? 77

3.2.2 Conflicts of Interest, Bias & Objectivity 80

Part II The Normative Background **87**

4 Imbalance and Trustworthiness **89**

4.1 A Definition of Trustworthiness 91

4.2 The Trustworthiness of Science 94

4.2.1 Requirements on the Public Epistemic Trustworthiness of Science 94

4.2.2 *x*: The Object of Trust & Trust in Objects 102

4.2.3 *A*: The Trusting Party—Who is "the Public"? 105

4.2.4 *B*: The Trusted Party—Scientists & "Science" as an Institution 109

4.3 Summary 113

5 Imbalance and Productiveness 117

5.1 Imbalance & Misleading Claims 118

5.1.1 Imbalance as Miscommunication 122

5.1.2 Imbalance as Alethic Risk 125

5.2 Towards the Framework 135

5.2.1 Challenges 135

5.2.2 The Criterion of Productiveness 146

5.2.3 The Alethic, the Practical and the Ethical 149

6 Imbalance and Injustice 153

6.1 Neglected Diseases & Justice 154

6.2 Imbalance as Epistemic Injustice 162

6.2.1 Distributive Epistemic Injustice 162

6.2.2 Participatory Epistemic Injustice 172

6.3 The Criterion of Epistemic Injustice 181

Part III Results 187

7 Explaining Imbalance in the SoR 189

7.1 Application of the Framework 189

7.1.1 Relations between the Criteria 191

7.1.2 Argument Templates 195

7.1.3 Application to the Examples 203

7.2 Discussion 225

7.2.1 Evaluation of the Results 225
7.2.2 Imbalance and Higher-Order Concepts 227
7.2.3 Outlook 233

Chapter 1

Introduction

In many contexts, when we want to express that some advice, examination or judgment is dependable, we might refer to such information as independent: if there is suspicion that some organization—be it a firm, a part of the government or maybe the church—suffers from bad practice, we call for an independent inquiry. To arrive at a well-founded opinion about current events, we may look towards the independent press. For any decision where different interest groups stand to benefit—or suffer—, we might want to seek independent information about the facts, which is supposed to give us some common ground from where to start our deliberation. At least one—if not *the*—major source for such independent facts about the expectable consequences of our actions and the world in general, is supposed to be scientific research.

Yet in recent debates, the idea of science as an independent arbiter has often been called into question: one major cause for concern is the ever-expanding entanglement of science with economic interests—think of the research sponsored and managed by the tobacco industry (cf. Gruening, Gilmore, and McKee 2006; Oreskes and Conway 2012), or, more recently, the sugar industry (cf. Kearns, Glantz, and Schmidt 2015). But doubts have also been raised about the political impartiality of science. The field of climate science, for example, has come under scrutiny concerning what has come to be called "Climate Gate"—where leaked emails between scientists "were cited by climate change critics as evidence that British and American scientists had changed their results to make global warming appear worse than it is" (Leiserowitz et al. 2013, 819), although subsequent investigations did not substantiate these claims. Greater independence, however, is not always what is demanded by critics of modern science. Daniel Sarewitz (2016), who has gone so far as proclaiming an impending "Scientific Doomsday" describes parts of modern research as too egocentric—focusing on the career interests of researchers and the inner logic of academia where it should be accountable to the needs of members of the public instead. But why should the independence of information be considered valuable in the first place? In the context of individual self-realization, autonomy or independence may be desired for its own sake. But in all the uses of "independence" mentioned above, the source of information being independent is not a goal in itself; rather, it is an instrumental value which is supposed to ensure some other quality: by being independent, the information provided is supposed to be free from undue influences, and thus dependable and useful for all those who may rely on it. In the context of science, there is one other concept connected to such ideas, which also has a long tradition within discussions in the philosophy of science: scientific objectivity.

When we think about how the objectivity of science can be threatened, we might first think of biased decisions made by individual scientists. The trinity of research misconduct—fabrication, falsification, and plagiarism—comes to mind. Over the last decades, many research institutions and professional associations have adopted guidelines to prevent these problems or at least make visible the conflicts of interest that—among other things—might trigger them (Pigman and Carmichael 1950; Whitbeck 1995; Lo and Field 2009). But individuals whose decisions influence particular research projects are only part of the problem at hand.

Articles in popular science magazines based on just one finding might promise coming innovations or ask us to change our lifestyle. In public discussions, individual scientists might appear as representatives of whole fields. But if we take the idea of science as a source of dependable knowledge that is supposed to inform our actions seriously, we should not just consider individual findings—or, perhaps even more troublesome, the opinions of individual scientists. Instead, as I will argue here, we should be informed by the entirety of what is considered state of the art in science: the multitude of currently considered scientific evidence, hypotheses, theories, and traditions, be they accepted or contested, in agreement or in disarray. For one thing, this diversity of research efforts might give us—in the vein of Longino's (1990) plea for a pluralistic idea of objectivity—the hope that the scientific community might, through avenues of criticism, overcome flaws on the individual level. But this perspective also gives rise to an emergent set of worries: not only might particular bits of research be false, or individual scientists be biased. There are particular challenges for the independence and thus the objectivity of science that only emerge at the level of what I will call the overall *state of research*: the distribution of research projects and findings on certain topics or with particular outcomes can be problematic even when all the individual results are fine. Considering the whole set of available—or at least, seriously considered—information poses the additional challenges of comprehensiveness and balance of scientific findings. But a conceptual framework for the discussion of such imbalances and their ethical and epistemological implications is sorely lacking. More specifically, while we have the conceptual resources to make sense of problems with the objectivity of individual results, we lack a parallel understanding of what objectivity is supposed to mean when it is applied to the entire distribution of research findings relevant for a topic or decision of interest. In this dissertation, I will show that there is more to this second challenge than just aggregating the individual findings' objectivity. Over the course of this dissertation, I will take up several problems with imbalance in the SoR that have been discussed individually in the literature—not only in philosophical discussions, but also in other meta-scientific disciplines and also comments by scientists themselves. Based on these examples, I provide a detailed systematic account of what constitutes imbalance in the SoR, and, in the end, explain what this implies for existing concepts such as independence and scientific objectivity.

1.1 Overview

The text of this dissertation contains three parts and seven chapters:

Part I—Imbalance in the State of Research (SoR): In the first part of the dissertation, I describe the phenomena under investigation and clarify the two main concepts: the state of research and its imbalance.

In **Chapter 2: What is the SoR?**, I explore the concept of "state of research".

There is a principal distinction to be made between the overall state of research in the abstract—understood as the entirety of facts about scientific research at a specific time—and the state of research connected to a specific topic, question, or decision to be made. I discuss different types of reconstructions that present such specific SoRs in different media both in science—think of meta-analyses— and beyond. With a focus on different types of evidence and evidence synthesis, I discuss some of the most important aspects of the SoR reported in such reconstructions. Towards the end of the chapter, I argue that—contrary to the broad idea of an overall state of research—the concept of a specific SoR necessarily has normative underpinnings: the scope of the SoR depends on the action-guiding argumentation that connects the SoR to policy decisions. This model of relevance introduced in the last section of the chapter will be important throughout the dissertation.

In **Chapter 3: What is Imbalance?**, I introduce the notion of imbalance in the state of research. Instead of taking a theory-driven approach, I provide a list of examples—a variety of different mechanisms and corresponding cases from the literature—which describe distributions of research which seem intuitively problematic. In a first attempt at categorization, I discuss similarities and differences between these cases. Taking up the notion of relevance from chapter 2, I go on to describe how the distribution of research interacts with decision-making further down the line. What distinguishes these cases from other issues with scientific research, I argue, is that they are about gaps or overabundance of types of research, not about the quality of individual results.

In the last part of this chapter, I argue that it is not trivial to distinguish imbalance from acceptable or even desirable forms of unequal research distribution; in this respect, imbalance is similar to connected value-laden notions such as independence, conflicts of interest, bias and objectivity. Thus, to explain cases of problematic imbalance, one needs to substantially engage with normative background theories, which I do in the second part of the dissertation.

Part II—The Normative Background: In the second part of the dissertation, I engage with existing discussions in philosophy of science, social epistemology, and research ethics to provide a normative framework for assessing the examples of imbalance in the state of research that I described in the first part.

In **Chapter 4: Imbalance and Trustworthiness**, I present (epistemic) trustworthiness as a normative concept that can be used to connect the different issues discussed in chapter 3. I present a general framework for trustworthiness and apply it to public (epistemic) trust in science. I focus on the requirement of scientific integrity, which connects trust with scientific freedom: science needs to be productive and fair to warrant the public's trust. These two features—epistemic productiveness and epistemic justice— lead to the next two chapters of the dissertation.

In **Chapter 5: Imbalance and Productiveness**, I discuss imbalance as distributions of research that do not reflect the objective of epistemic productiveness, i.e., efficiently producing knowledge which is useful for society. I begin by exploring the idea of imbalance as an alethic risk—the risk that someone might hold false beliefs as a consequence of the available research –, distinguishing issues of miscommunication and more substantial forms of bias. I then argue that alethic risk alone is not enough to explain why, even from the perspective of epistemic risk, imbalance is a problem: if science is supposed to be useful in a democratic society, it has to avoid not only the risk of false beliefs but also of errors in practical decision-making.

In **Chapter 6: Imbalance and injustice**, I turn to the second criterion demanded by the public trustworthiness of science: epistemic justice. In the first part of the chapter, I argue that for some of the examples which have been interpreted in terms of productiveness, it would be more appropriate to consider an explanation in terms of fairness: rather than a problem with maximization of the SoR's contribution to society, inequalities concerning the benefits to various groups of people are at the heart of these problems.

In the second part of the chapter, I take up this issue by discussing imbalance from the perspective of a broad conception of epistemic (in-)justice: not only the consequences of certain distributions of research simpliciter but the question of who is affected by these consequences and who is not, both positively and negatively, can matter for assessing imbalance. In addition to this idea of epistemic distributive justice, epistemic injustice also allows us to consider cases where not only the consequences but already the access of different groups to the decision-making process which governs the distribution—e.g., decisions about funding, pursuit, and publication—is unjust.

Part III—Results: The third and last part of the dissertation consists of one more chapter in which I combine the discussions from the previous parts.

In **Chapter 7: Explaining Imbalance in the SoR**, I first discuss the relations between the criteria introduced in the previous chapters and then apply this framework to the examples of imbalance from chapter 3. Lastly, I evaluate the results of this application process and the implications for the other concepts related to imbalance: independence, conflicts of interest, bias, and objectivity.

In the following section, I explain the rationale behind this structure by relating it to my main research interests and describing my general approach.

1.2 Research Questions & Approach

What is scientific objectivity? This very general question looms behind the discussion in the first two sections of this introduction. More specifically, I am interested in the kind of objectivity we may expect from independent sources of information; from epistemic products of scientific research, where these are valued not just in themselves, but as a resource for policy-making and important individual decisions. That is, I do not aim at providing a universal account of scientific objectivity in general. This goal would also not be very fruitful, because, as authors like Heather Douglas (2004) have argued, we have reason to believe that objectivity is not one monolithic value, but rather a nexus of multiple irreducible concepts, answering to more specific contexts of application. This brings me to the following background question:

> **Background question:** What is the meaning of scientific independence and objectivity in the context of practical decision-making?

For one, this means that I am interested in a concept of scientific objectivity with practical relevance; also, it implies that it has to be applicable not only to one finding or one theory in isolation, but to the entire base of information needed to make these practical judgments—at least as long as the sciences can provide them. Certainly, also the objectivity of an individual result will have some influence on the overall objectivity. The objective of my research, however, is to explore the question of what new challenges for objectivity emerge when we leave aside the problems with isolated claims about the world. In this dissertation, I will show that the objectivity of an entire research landscape, of the state of research on a certain topic, is more than that. Rather than about "getting at objects in the world" (ibid., 472), the objectivity of states of research is about providing a balanced account of relevant matters. This leads me to my main research question, the answer of which is supposed to provide insights into the broad background question stated above:

> **Main research question:** What problems emerge when we focus on the overall distribution of research, rather than individual scientific findings?

That is: What is imbalance in the state of research?

First, there is some theoretical work to be done. "State of research" is not a well-defined term, neither within philosophy of science nor in other fields. A clear concept of state of research, however, is needed to guide the discussion in the rest of the project. Thus follows a first subordinate research question:

> **Subordinate question 1:** What is a state of research?

My interest in states of research is not just motivated by a theoretical understanding of a particular level of aggregated knowledge. It is connected to the overarching theme in which scientific objectivity and practical decisions are brought together. Working towards an adequate definition, I therefore analyze specific forms of scientific literature that aim at providing reconstructions of the state of research for particular audiences: most scientific articles contain a theory section, trying to summarize the field for interested readers. Review articles or meta-studies more directly try to establish something like a state of research concerning specific topics. Even more geared towards application, policy reports summarize scientific findings for politicians.

Explaining the connection between the state of research and its application also requires a concept of scope, which makes it possible to determine if some research project is relevant for policy or individual decision-making. The consequences of imbalance, however, are not restricted to its effects on applications outside of science. As Elliott and McKaughan (2009) show, scientific methods and the appraisal of hypotheses and theories are affected by what candidates are available—and thus, by the balance of state of research. With the idea of "value-laden blind spots", Anke Bueter (2015) describes a very similar mechanism. With both epistemic and practical relevance in mind, one has to ask:

Subordinate question 2: When is a product of scientific research relevant for a topic of interest or a decision to be informed?

Besides the state of research, there is another opaque term contained in the main research question: what does it mean for a state of research to be in "balance"? I will begin with an analysis of particular cases of imbalance in the SoR; that is, I explore examples where there seems to be a problem with the distribution of scientific research. I thereby follow Miranda Fricker (2017), who—in the context of justice—has provided a general argument for starting from the individual and dysfunctional, rather than ideal concepts:

> The interest in the dysfunctional and the non-ideal need not stem from an intrinsic interest in these things (though they are indeed interesting); rather it may stem simply from a realistic interest in how to achieve functionality in any given practice. Thus a philosopher who only aimed to understand and represent epistemic practices in their most functional forms, even in some notionally ideal form, would still need to do so by looking first at what potential collapses into dysfunctionality are being perpetually staved off, and by what mechanisms. (ibid., 57)

By focusing on what can go wrong with the state of research, I hope to improve our understanding of what criteria should be fulfilled for it to count as balanced, and thus also, to our understanding of higher-order concepts such as independence and objectivity. To do so, however, one first needs to provide support for a positive answer to another question:

Subordinate question 3: Are there cases which constitute problems with the overall distribution of research, rather than individual results?

It should not come as a surprise that I will claim that there are; the important part of the answer, therefore, is not just that they exist, but what those examples are, and how they are described. What first got me interested in the state of research and distributions of research rather than the justification of individual claims were isolated examples such as the idea of unpatentable research: James Robert Brown (2008) argues that the influence of private sponsors in medical research, even where it does not distort the outcomes of studies, may still be problematic because it restricts the research agenda to those projects which may lead to patentable and thus possibly profitable products, such as medical drugs—while alternatives, such as sports, diets or social factors, remain neglected. Starting with a few initial examples, throughout the project I collected a variety of cases that supplement each other by providing novel aspects to the discussion. In the first part of the dissertation, I present this wide range of cases that are connected by the common theme that there is too much or too little research of a particular type. Examples stem both from philosophy of science (Flory and Kitcher 2004; Brown 2008; Sismondo 2008; Stegenga 2011), the wider meta-scientific literature (Sismondo 2008; Frickel et al. 2010; Song et al. 2010; Chan et al. 2014; Lewandowsky, Risbey, and Oreskes 2016) and also from the comments of concerned scientists themselves (Edwards et al. 2011). At this point, it is important to stress that this is a strictly philosophical dissertation. That is, while I will engage with, for example, psychological, sociological, or historical examples to ground my arguments and illustrate my claims, my own contribution will mostly consist of conceptual and normative work. A substantial empirical inquiry is beyond the scope of the project. Still, a thorough conceptual analysis of imbalances in the state of research will not only contribute to philosophy of science but also help systematize the phenomena and thus allow for an improved theoretical grounding of future research questions, connect various findings, and point towards less explored phenomena, inspiring further research.

The list of examples by itself, however, does satisfy neither goal; it still provides no satisfying answer to the question of what imbalance is, of what makes it problematic. I will begin by discussing the systematic similarities and differences between the individual examples. Then, I will ask what exactly is problematic about each of them. This leads me to one further subordinate research question, which will be of importance throughout the second part of the dissertation:

> **Subordinate question 4:** How can we explain why the examples of imbalance should be considered problematic?

First, I will try to explain the intuitive judgments contained in each of the case descriptions by referring to some prolific concepts which are used in the context of problems with science as a source of information, such as conflicts of interest, bias, or independence and objectivity themselves. All of them, or so I will argue, are either too vague or too ambiguous to provide clear criteria that can help explain why one could judge the examples problematic. In the second part of the dissertation, I therefore systematically develop a normative framework that can accommodate the list of examples. Starting with the concept of epistemic trustworthiness, I consider several normative principles taken from the philosophical discussion, and discuss how they might be applied to the examples. In a process of mutual adjustment, I made changes both to these systematic normative underpinnings and the analysis of the individual cases until the final framework emerged. In the third part of the dissertation, I evaluate this framework by using the criteria contained to provide one or more explanations for each of the examples, showing that is both comprehensive and coherent. Then, I turn back to the broad question in the background of the project, and answer one final question:

> **Subordinate question 5:** What are the implications of the concept of imbalance in the SoR for higher-order concepts such as independence and objectivity?

I will argue that, on the level of the SoR, both independence and objectivity have to be about more than just "Faithfulness to the Facts" (Reiss and Sprenger 2017). Scientific research, as a source of information for the public, has to aim not only at isolated truths, but also at providing practical benefits, and do so in a manner that is fair to the interests of the various social groups that rely on science.

Part I

Imbalance in the State of Research (SoR)

Chapter 2

What is the SoR?

In the introduction, I have claimed that if we care about the usefulness of scientific research in decision-making, we should look towards the state of research (SoR) as opposed to just individual pieces of evidence and the claims they can support or challenge (chapter 1). In the rest of this dissertation, I will turn towards the concept of imbalance in the SoR (chapter 3), which is supposed to characterize SoRs that can be considered problematic for a variety of epistemological or ethical reasons (chapter 4 - chapter 6). Before I can start to explicate this notion of imbalance, however, I will have to define its subject—the SoR[1]—, which is the objective of the chapter at hand.

Very generally, we might say that the SoR is nothing but a particular facet of the state of the world, referring to the state of all of science; to all that could possibly be known about scientific research at a particular point in time. We can thus define a first, abstract notion of an overall SoR:

> **Def.: The overall state of research** is the set of all facts about scientific research at a given time.

"Fact", here, is not supposed to refer to true propositions, but to some aspect of the world, that, in principle, someone could know *about*. But what do I mean by "scientific research"? In a very general sense, it should be taken to mean any

1. "SoR" is not a technical term in the scientific, meta-scientific or philosophical discussion; there are other, similar terms—such as state of the science, state of the art, or others—which could be substituted here, although they might have slightly different connotations.

kind of scientific activity connected to the aim of producing novel knowledge about the world, increasing our understanding or correcting our existing beliefs. In this dissertation, however, I will mostly focus on research in terms of written reports of those investigations, as they are easier to access than other forms of research, and also because, for the most part, SoRs used in decision-making are based on such material.

There is an argument to be made that the set of facts about scientific research is infinite. For example, we may grant that the set of facts about the research also includes the set of facts about how it was produced—and thus about its history. But for any event, such as the production of some piece of research, there exist innumerable causes, going back, at least in principle, to the beginning of time. We can consider any cause in the history of the research at any point, resulting in an infinite set of facts. Philip Kitcher makes a related point about facts about the cosmos in general (cf. Kitcher (2011, 106)). Like him, we might want to say that although there is this multitude of facts, only a small fraction of them will be interesting or, to use Kitcher's term, significant.

This gives rise to a possible objection against the definition given above: the SoR—or so one might argue—is supposed to be informative, only containing interesting facts for the topic at hand or the decision to be made; it is supposed to report the state of the art in science, give us the cutting edge knowledge needed to make the decisions we are interested in. The definition of SoRs given before, therefore, would be much too inclusive: not only does it apply to an infinite amount of facts—which is more than we can ever hope to process—many of these facts will also simply not be useful to us. This objection, however, rests on an ambiguity of the term "SoR": sometimes we refer to the SoR in the abstract, as a part of state of the world—but sometimes we also refer to specific perspectives on this facet of reality, motivated by a particular interest in scientific research. My initial definition captures the former; the latter refers to the subset of facts about the research which are important from that perspective, and which can be reported in specific accounts of the SoR. Keeping the definition above broad as it is, allows for different forms of such accounts, focusing on different aspects of the SoR.

The need to keep a concept of SoR which is open to various perspectives is emphasized by the historical change concerning what may be considered the important facts about the research: consider feminist philosophers like Helen Longino (2008), who points towards the influence of certain values held by different social groups in the research community as an important feature of the SoR in areas like primatology. Before the advent of social epistemology, considerations of the social background of researchers would probably not have been considered an important or even admissible fact in accounts of the SoR. The wide definition of a SoR in itself thus also allows us to keep an open mind for further important aspects of the SoR we have not yet considered.

However, the alternative concept of a specific SoR connected to a particular topic will be the one used throughout this dissertation. After all, I am concerned with the SoR as a particular level of knowledge, which can be used as a resource for making specific decisions. In section 2.1, I will approach this second notion of SoR by examining what accounts of the SoR can actually look like; I will give examples for different media in which such accounts are presented, and consider possible audiences and the function of presenting the SoR in different ways. In section 2.2, I will point towards different types of facts that often are included in accounts of the SoR: For each specific account, one needs to decide which aspects of the SoR to report, and I will highlight some of the most important ones, focusing on different types of evidence. Finally, in section 2.3, I will address the question of relevance or scope of SoRs. There I will argue that even on a conceptual level, we cannot determine what the SoR related to a decision or topic is without committing to a specific normative perspective.

2.1 The SoR and its Reconstruction

We might be interested in the SoR for any number of reasons. But whenever we want to make use of facts about some area of research, someone will first have to investigate and prepare the data in a way that fits the purpose at hand. Sometimes the recipients of the information will be able to do this themselves—for example, when scientists try to find out about the current state of their field before they decide which projects to pursue. But sometimes, we also have to rely on the accounts provided by others, be it for lack of expertise or opportunity. I have chosen the term “reconstruction” to describe all of these practices that provide an account of the SoR.

There would have been several possible alternatives, such as "description" or "representation". But the notion of reconstruction has a particular connotation, highlighting the active role of whoever provides the account. We should not be tempted to think of the process of providing information about the SoR as an objective enterprise in the sense that there is only one right way to decide what to include and how to present it. The relationship between the SoR in itself and its reconstruction is not one of approximation. There is no one ideal reconstruction of the SoR, which—akin perhaps to the "ideal explanations" of Peter Railton (1981)—would report all the infinite facts about the research topic of interest. Instead, reconstructions of the SoR are to be seen in the context of their use. That is, reconstructions should not report as many facts as possible in the most detailed way; instead, their form and content should follow the intended function. When we are interested in the SoR, we often simply will not be interested in all the facts. What we are after might be an overview of the current best knowledge available in science. This concept of SoR-reconstruction presupposes a normatively laden selection process, where some research is discarded as obsolete and replaced by a better understanding of the subject matter. For example, when we are interested in the SoR on a particular matter in mechanics as a subfield of physics, it would seem intuitively unfitting to list, among textbook physics, medieval texts on Aristotelian kinematics. Also, for reasons of speed of production or ease of understanding, it will often be advisable to include less rather than more information in a reconstruction. This is not to say that it would not in principle be possible to try and provide a maximally comprehensive reconstruction of a SoR—i.e., one that reports as many facts about the SoR as possible—just as it is possible to go more and more into detail about any subject. However, this is neither to be considered the proper form of reconstruction nor even a very typical one. We can now give a definition, differentiating the SoR and its reconstruction:

> **Def.: A reconstruction of the SoR** is an account that aims at conveying particular aspects of the SoR, selectively reporting a subset of all facts about the research on a topic at a given time.

Accounts that can be considered reconstructions of SoRs can occur in various media and formats. In the scientific literature, they appear in the form of scientific reviews. This refers not only to entire review articles but also to reviews which appear as parts of larger publications. The genre of scientific review is anything but homogeneous. Grant and Booth (2009) distinguish fourteen types of reviews, which are presented with their methodology and—as the authors put it—different strengths and weaknesses. While there is considerable overlap between the types laid out in the article, it also quite clear that there is a wide spectrum of reviews, some of which are more suited to certain purposes than others. In the following, I will examine two examples for reconstructions of the SoR: the Cochrane Review, a form of systematic review, and literature reviews as they appear in the introduction section of primary research articles. These exemplary types of review are very different both in form and function, which will help to illustrate how one follows from the other, while also providing a more detailed picture of how reconstructions and the underlying SoR are connected. However, these are only two samples from the spectrum of reviews and do not by any means cover the many different ways in which the SoR is reconstructed within science.

The Cochrane review. Cochrane, an international NGO known for their support of the "evidence-based medicine"-movement (EBM) champions the Cochrane review, a form of systematic review. They also maintain the Cochrane Library, an online database in which these reviews are collected. It is the self-proclaimed aim of Cochrane to support the well-informed decision making of "healthcare providers, consumers, researchers, and policy makers" (Higgins and Green 2008, 6). The systematic reviews are supposed to provide decision-makers with evidence concerning very specific research questions that can inform "practical decisions about health care"[2] (ibid., 13).

> A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing

2. Note that, in the following, I only discuss ideas of how different types of reviews are intended to fulfill certain functions; I do not want to make more substantial claims about if they do or do not actually succeed.

> bias, thus providing more reliable findings from which conclusions can be drawn and decisions made [...]. (Higgins and Green 2008, 6)

In Cochrane reviews, eligibility is interpreted in terms of the criteria given by the "acronym PICO (Participants, Interventions, Comparisons and Outcomes)" (ibid., 84). This results in only very specific studies—which are sufficiently similar in all of these categories—to be selected after an initial keyword search in relevant medical databases. Narrowing the scope of research under review in this way is supposed to make the studies included comparable—i.e., results of multiple studies can be aggregated—while safeguarding the reliability of the results. In contrast with other types of reviews, this process of selecting the material under review is supposed to be transparent and reproducible, thus avoiding selection bias (ibid., 97).

The systematic methods mentioned in the quote above also include methods for statistical analysis, which are supposed to provide further transparency and reliability. The method most widely used in Cochrane reviews is meta-analysis, a tool for combining the results of multiple separate studies. Jacob Stegenga (2018) explains the underlying rationale:

> In contrast with qualitative literature reviews and consensus conferences, meta-analyses have a constrained structure and a quantitative output. The importance of using systematic methods of amalgamating evidence became apparent by the 1970s, when scientists began to review a plethora of evidence with what some took to be personal idiosyncrasies. (ibid., 85)

The constrained structure of the Cochrane review, which is supplemented by the comprehensive guidelines contained in the *Cochrane Handbook for Systematic Interventions* (Higgins and Green 2008) and the statistical methods, can be thought of as an attempt to achieve procedural objectivity in the vein of Douglas (2004). The quantitative methods combining the evidence from individual studies also promise "an increase in power, an improvement in precision, the ability to answer questions not posed by individual studies, and the opportunity to settle controversies arising from conflicting claims" (Higgins and Green 2008, 242).[3] Epistemic virtues such as

3. Stegenga, however, criticizes the EBM movement for their focus on meta-analysis, claiming

objectivity, reliability, and precision are, of course, sought after in many contexts in science. But the clear focus on providing aggregated scientific results which can be used in practical decision-making concerned with medical interventions—where small mistakes can have very serious consequences—gives a clear pragmatic-epistemic background to Cochrane reviews, affecting which aspects of the SoR are considered and how they are presented.

Introductory reviews. A second example points in a different direction: since the second half of the 20th century, most primary research articles—at least in the medical sciences—follow what is sometimes called the IMRAD-structure: they include the standardized sections of introduction, methods, results, and discussion (cf. Sollaci and Pereira 2004). The introduction part contains a reconstruction of the SoR—which I will refer to as "introductory reviews" for lack of an official term:

> The first paragraph should be a short story of the current knowledge of the attempted research area (to state "what we know" of the problem that was investigated). This should lead directly into the next paragraph that summarises what other people have done in that field, what limitations have been encountered to date, and what questions still need to be answered (to speculate "what we don't know"). (Todorović 2003, 203)

Note some interesting aspects of this description: while the systematic review aimed at eliminating idiosyncrasies of the reviewer, the quote above invites them to "speculate", and provide a "story", i.e., a narrative text. Also, it is not—or at least, not only—supposed to collate existing evidence, but also to give an account of what is not available but perhaps *should* be—implying a normative judgment. In contrast with Cochrane reviews, introductory reviews are not supposed to include very specific material, nor do they require a comprehensive search. On the contrary:

> However, in this section, one wouldn't review all the literature available. One must resist the temptation to impress readers by summarising everything that has gone before. They will be bored, not impressed, and will probably never make it through the present study [...]. (ibid., 203)

---

that it does not avoid malleability to a sufficient degree, and thus the promised constraint is often not achieved (cf. Stegenga 2018, chapter 6).

Why do introductory reviews take this form? A first answer is that, of course, being a small part of a research paper that has to accommodate the space restrictions of journal publications, they cannot be much more comprehensive. Another answer is that, once again, this type of review is supposed to fulfill a particular function, or rather, a plurality of functions: an introductory review 1) "tells why the reader should find the paper interesting", 2) "explains why the author carried out the research", and 3) "gives the background the reader needs to understand and judge the paper "(Nair and Nair 2014, 18). A last reason which might also apply to introductory reviews is given by Grant and Booth (2009) in the context of the type of "narrative review": 4) "to demonstrate that the writer has extensively researched the literature and critically evaluated its quality" (ibid., 93).

The first two reasons come down to arguing that the research paper is important in the context of a given discussion; this necessitates the kind of normative judgment about lacunas in the SoR highlighted above. It also explains why this type of review does not necessarily have to be comprehensive: while claiming that there is no or insufficient research into some research question does require the reviewer to have a complete picture of what has been done, it is neither fruitful nor possible to provide the readers with a complete list of research which does not respond to the question at hand. It suffices to argue that there is existing research which implies a further question and that an investigation of it is pending as of now.

The third function of an introductory review mentioned here—giving the reader the background he needs to understand and judge the paper—does not require an account of all preceding research either. It only implies that the concepts used in the paper need to be referenced, not that the reader is acquainted with all research on the topic in detail. One could argue that the information about the available research could also be used by readers as a reference for conducting their own literature surveys on the topic, which might profit from a more comprehensive overview. But firstly, this is—at least according to the authors referenced above—not the primary function of introductory reviews and might be better served by dedicated literature review articles. Secondly, the need to be very selective of references could even be helpful for such survey purposes, as the selection could be thought of as an indicator for which previous work the author of the research paper considered especially important or relevant for the discussion in the discipline.

This ability to select certain important studies and connect them in the narrative can also be thought of as a way to address the fourth and last reason given above: this signals to potential gatekeepers in the publication process that the author has a sufficient understanding of the scientific debate. In contrast, comprehensiveness was definitely desirable in the case of the systematic review, where the objective was to provide evidence in aggregated form, and where failing to include all relevant studies would bias the outcome of, e.g., meta-analysis.[4]

Not only do these functions differ from the systematic review, but the introductory review is generally much less focused. While it can be argued that the third objective above has some direct epistemic merit for the readers, there are clearly also non-epistemic, pragmatic considerations involved, such as establishing the significance of the research and the expertise of the author. The systematic review, on the other hand, seems to have a clear epistemic goal: providing an evidence synthesis for a very specific issue while avoiding bias on the side of the reviewer.

There are many more examples for reconstructions, also apart from scientific reviews. Still in academia, but outside the realm of research, we find textbooks referencing the SoR, which are aimed at educating students and teaching. Similarly, scientific handbooks introduce interested parties to specific topics. Another big sector which produces reconstructions aimed at a non-scientific audience—or at least at readers from other disciplines—is science journalism. Once again, this field imposes many restrictions on form and content: the time spent on an individual article might be severely limited compared to scientific reviews; the language and depth of the texts must be appropriate for the target audience; and the markets which the including publications cater to might be quite different from academic publishing. That is not to say that the latter is not heavily influenced by economic considerations. However, while the primary currencies for scientific publications are citations and journal impact, popular science magazines have to follow the more general rules of a journalistic sales

4. We will revisit this point in section 2.3, where I discuss Stegenga's criticism of meta-analysis as the platinum standard in medical research. It will also make a reappearance in the example of the evidential standards of EBM, which is used throughout the later chapters of the dissertation.

market.

Reconstructions for policy. I cannot give a detailed analysis of all of these contexts, the functions for reconstructions they bring with them, and what this implies for their form and content. But as the objective in this book is to shed some light on how the SoR and its reconstructions affect decisions, especially in the context of policy-making, there is one last area that should get some special consideration: reconstructions of the SoR in policy-advice. The aim of informing the practical decision-making of policy-makers was also given for Cochrane reviews; compared to other forms of scientific policy advice, however, they constitute somewhat of an outlier. The topic of Cochrane reviews are not decision-situations; they are attached to very specific research hypotheses, and they are structured according to scientific inferences about the confirmation of these claims.[5] This is not to be confused with the claim that systematic reviews do not give recommendations, while policy advice needs to do so. While Higgins and Green (2008) do indeed explicitly state that "[a]uthors of Cochrane reviews should not make recommendations" (cf. ibid., 380), the same is true of paradigmatic policy advice reports such as the IPCC assessment reports as well:

> As with all IPCC products, the report is the result of an assessment process designed to highlight both big-picture messages and key details, to integrate knowledge from diverse disciplines, to evaluate the strength of evidence underlying findings, and to identify topics where understanding is incomplete. The focus of the assessment is providing information to support good decisions by stakeholders at all levels. The assessment is a unique source of background for decision support, while scrupulously avoiding advocacy for particular policy options. (Field et al. 2014, ix)

In fact, to what degree scientists as policy-advisors should endorse specific courses of action is an issue of lively debate in science itself but also in the meta-scientific disciplines. Political scientist Roger A. Pielke Jr. 2007 distinguishes four ideal types of policy advisor: the "pure scientists", the "science arbiter", the "issue advocate" and the

5. I follow up on the difference between these two levels in section 2.3.

"broker of policy alternatives" (Pielke 2007, 1-2). But no matter where a reconstruction of the SoR falls in this scheme, what makes it a form of policy advice is that its content is dependent on the policy options in the decision-situation with which policy-makers are faced: for example, while certain scientific hypotheses such as the existence of anthropogenic climate change are a central issue in the IPCC reports, one just has to consider the self-description in the IPCC Fifth Assessment Report:

> Topic 4 (Adaptation and Mitigation) describes individual adaptation and mitigation options and policy approaches. It also addresses integrated responses that link mitigation and adaptation with other societal objectives. (Core Writing Team, Pachauri, and Meyer 2015, 36)

It is evident here that the underlying policy-options determine the scope of the reports. While we could also consider Cochrane reviews to be decision-relevant in that they give answer to very specific questions about decisions in healthcare, this broader, policy-oriented view is lacking. Cochrane offers the format of "Cochrane Overviews" for such purposes, where:

> [...] a central aim [...] is to serve as a 'friendly front end' to The Cochrane Library, allowing the reader a quick overview (and an exhaustive list) of Cochrane Intervention reviews relevant to a specific decision. The primary audiences envisioned are decision makers (such as a clinicians, policy makers, or informed consumers) who are accessing The Cochrane Library for evidence on a specific problem. (Higgins and Green 2008, 608)

In section 2.3, I will return to the question of what provides the criterion of relevance for SoRs in the context of decision-making. But first, I will give an overview of different aspects of SoRs that can be included in their reconstruction: in the current section, I have argued that different purposes for reconstructing SoRs give rise to a multitude of forms for such reconstructions. But another important aspect is what kind of content we expect when we ask for the SoR on a particular topic. With the difference between introductory reviews—which are supposed to contain a non-exhaustive list of what research has done on a topic, and what is still lacking—and the systematic review—which provides aggregated results for a specific research question—we have already encountered different types of facts demanded by different purposes. In the next section, I will expand this list, most importantly by giving an account of different types of facts that can be considered evidence.

2.2 Aspects of SoRs

The SoR consists of all facts about the research on a specific topic; however, as I have argued above, when we talk about SoRs, we mostly do so with a specific goal in mind—such as informing decision-making—which determines what aspects of the SoR are of interest to us. As we have seen in the last section, the way in which the SoR is supposed to help decision-making is to provide the decision-makers with an overview of the available scientific evidence. Douglas (2012) asks us to "consider all the available relevant scientific evidence when making policy-decisions, whether those decisions concern the extent to which a policy is needed or the exact nature of the policy intervention" (ibid., 140); in a history of research synthesis, Chalmers, Hedges, and Cooper (2002) describe its prominence as arising from "a need to organize and evaluate the accumulating bodies of research evidence" (ibid., 19); Grant and Booth (2009) similarly argue that it "quickly became apparent that synthesized summaries of 'all' evidence within a particular domain would be required, in addition to the evidence from primary studies, if clinicians were to make truly informed decisions [...] (ibid., 91)."

Evidence is indeed perhaps the primary aspect that would be reported about the research when asking for the SoR on some topic—but what exactly is evidence and what types of facts does it correspond to? First, I should note that no particular type of thing simply *is* evidence on its own; being evidence is a relational predicate, that is, to say that x is evidence is to say that x is evidence *for* some y. What exactly characterizes this relationship between x and y is subject to philosophical debate: Thomas Kelly (2016) differentiates some major lines of thoughts such as the idea of evidence as "the kind of thing which can make a difference to what one is justified in believing"—x as justification for belief y—, "evidence in the sense of reliable indicator"—x as an indicator for y—or the idea that if "E is evidence for some hypothesis H, then E makes it more likely that H is true"—x confirming y. In section 2.3, I will define the relevance of facts in the SoR for decision-making, referring to both the structure of scientific inference and, on another level, of moral argumentation. In this context, I choose to adopt the first, justificatory notion of evidence, which is closest to this argumentative approach. But what does the concept of evidence as justification for belief tell us about types of facts in the SoR and its reconstructions?

Consider an example from climate science: Petit et al. (1999) is a primary research article relying on evidence from ice-cores provided by drilling operations above Lake Vostok in Antarctica. This ice from deep under the surface was affected by the climate in the distant past, allowing for inferences to these conditions—the "Vostok ice-core record". For example, the results of an analysis of the composition of gas enclosed in bubbles within the ice at different depths were used to justify a data-set representing concentrations of CO_2 and CH_4 in the atmosphere at various points in history. Another important feature of the ice-cores is the ratio of different deuterium (δD) isotopes—and, alternatively oxygen isotopes (δ^{18}O)—in entrapped hydrogen and oxygen, gathered both from new gas-samples and from previous research. Comparing these ratios with today's values, this data is used to justify conclusions about historic atmospheric temperatures. The authors conclude that, while most of the historical climate variability—e.g., temperatures and glacial ice volume inferred from the isotope ratios—can be attributed to orbital forcing—that is "to that of the precession, obliquity and eccentricity of the Earth's orbit" (ibid., 429)—their evidence about CO_2 and CH_4 in the atmosphere compared with the temperature data suggests that these gases constitute an amplifying factor and "have contributed significantly to the glacial–interglacial change" (ibid., 435).

This argument is taken up Masson-Delmotte et al. (2013)—the chapter of the fifth assessment report of IPCC working group 1 that deals with paleoclimatic evidence: they report a large data-set about "[o]rbital parameters and proxy records" (ibid., 400), which combines the CO_2 and temperature data from Petit et al. (1999) with results from other research projects both on these and other parameters. This is then used as evidence in justifying the claim that "There is high confidence that changes in atmospheric CO_2 concentration play an important role in glacial–interglacial cycles", i.e., "temperature and ice volume changes" (Masson-Delmotte et al. 2013, 385), which is reported in the executive summary.

In these examples, we have already encountered two uses of evidence, which include different types of facts: 1) in Petit et al. (1999), the data about gas concentrations in the ice is used two justify beliefs about the historical concentrations of these gases in the atmosphere—facts about the objects of research (x) are used to justify another claim about the objects of research (y). These intermediate results are then used to justify their conclusion about the role of greenhouse gases in climate change, another factual claim of this type. 2) In Masson-Delmotte et al. (2013), the authors use results provided by other researchers concerning climate parameters to justify a similar conclusion—the hypothesis about the role of CO_2 in glacial-interglacial cycles. In this case, the facts used as evidence are not directly about the objects of research; instead, facts about the results of previous research (x) are used to justify belief in a fact about the research objects (y). Note that this second use of evidence—in which to some extent, researchers defer to the authority of other experts—is not exclusive to reconstructions of the SoR; also the primary research paper—Petit et al. (1999)—refers to previous results, e.g. concerning some isotope data (cf. ibid., 430) and the orbital forcing (cf. ibid., 431).

But we do not always have to take research results at face-value, and often we might not want to rely on them blindly. Consider a second example from the climate change literature: there are many papers that can be considered reconstructions of the SoR on global warming and that analyze the consensus among climate experts on the hypothesis that anthropogenic global warming does exist (Oreskes 2004; Anderegg 2010; Cook et al. 2016). The fact that this consensus exists, however, does not directly justify belief in the hypothesis; instead, it is used to establish the reliability of using some research results:

> In many senses, R10 [—Rosenberg et al. (2010), a consensus survey—] vindicates both the conclusions and the process of the IPCC. If the study had found large discrepancies between surveyed US climate scientists' views and IPCC conclusions, this could indicate a potential selection bias in the survey's selection or in IPCC author selection. However, R10's results in fact suggest that the IPCC's conclusions accurately reflect those of the US climate science community and

that bolsters confidence in the IPCC assessment process. (Anderegg 2010, 332)

In this example, the consensus in the scientific community is used to address concerns about the uses of evidence in the IPCC assessment. It thus indirectly contributes to the justification of belief in their results, introducing a third use of evidence: 3) facts about the evidence (x)—in this case, the agreement on its soundness in the scientific community—are used to justify belief in facts about the research objects (y). Expert consensus is not the only evidence providing such second-order reasons for belief: once again returning to the *Cochrane Handbook for Systematic Interventions*, authors of Cochrane Reviews are advised to perform "risk of bias" assessments which include the consideration of various facts about the included studies, such as if and how the assignment to experimental groups was concealed, if and how blinding was handled, and others (Higgins and Green 2008, 198-202). While these assessments then could theoretically be used to assign different weights to the results of different studies in eventual meta-analyses, currently, authors are instructed to only include studies with low risk of bias in meta-analyses, because appropriate methods are "not sufficiently developed" (ibid., 209).

Lastly, scientific claims about the research objects which so far have always appeared as the final target of justification (y), can be used as evidence for something else. The claim about the role of CO_2 in glacial-interglacial cycles in Masson-Delmotte et al. (2013), for example, could be taken up by policy-makers to justify a climate response strategy which aims at reducing the concentration of CO_2 in the atmosphere: if higher CO_2 concentrations can amplify climatic changes such as increasing temperatures, melting glaciers or sea-level rise, and we consider these events or their consequences harmful, reduction of CO_2 emissions or techniques such as carbon capture might be advisable forms of mitigation. In such an action-guiding—that is, perhaps, moral or political—argument, 4) the scientific claim, expressing a fact about the research objects (x) is used to justify a specific course of action, i.e., a decision option (y)[6]. This use of evidence was already apparent in the discussion of

6. I here interpret scientific claims as representing facts about research objects because—while they could also be thought of as facts about research results—usually they will enter action-guiding argumentation by justifying a belief about the world. In subsection 5.1.2, however, I will also suggest

the Cochrane review, where the specific research hypotheses that are addressed in these reconstructions were supposed to inform decisions concerning healthcare (cf. section 2.1). Note that the scientific claims discussed in reconstructions of the SoR may or may not be identical with the claims made in previous research. Systematic reviews that contain meta-analyses, for example, rely on published material, but by design are supposed to provide novel conclusions which are usually presented in forest plots. In a sense, however, even if reconstructions are purely cumulative, the selection and presentation of previous results does not only report the SoR, but add to it: they are re*constructions*, and constitute genuine—if secondary—research.

We can now give an overview of the four uses of evidence considered:

Types of Facts and Uses of Scientific Evidence

$\mathbf{E_1xy}$ Facts about research objects (x) used to justify belief in other facts about research objects (y).

$\mathbf{E_2xy}$ Facts about research results (x) used to justify belief in facts about research objects (y).

$\mathbf{E_3xy}$ Facts about other evidence (x) used to justify belief in facts about research objects (y).

$\mathbf{E_4xy}$ Facts about research objects (x) used to justify decision options (y).

These uses of evidence are connected (cf. also Figure 2.1):

In E_1 – E_3, evidence is used to justify facts about research in the form of scientific claims; it contributes to scientific inferences. The different types of facts mentioned here are not strictly mutually exclusive. Reconstructions of the SoR, at least in the form of scientific reviews, are to be considered secondary research; that is, they are research about research. In this type of investigation, research results and evidence are objects of the research; that is, in this context, both E_2 and E_3 appear as subsets of E_1. However, only those facts about research objects where these objects are themselves

that when evidence assessment is formalized, facts about the results might directly be used to justify a course of action; for example, where a certain number of studies affirming the effectiveness of some drug is needed for it to be approved by the FDA.

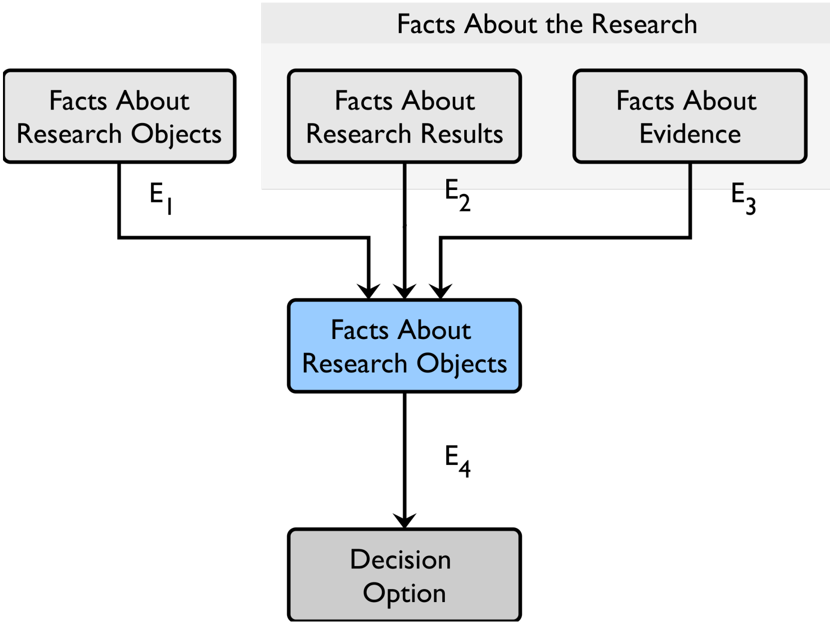

Figure 2.1: Uses of Scientific Evidence

research are facts about the research, and thus, per definition, part of the SoR. This means while primary research articles could in principle provide all kinds of evidence (x) appearing in E_1 – E_3, reconstructions of the SoR will only contain facts about research objects as evidence in the case they are also facts about other research, that is, as used in E_2 and E_3.

As I have argued above, E_4 is a very different use of evidence, in that it does not concern scientific inference, but action-guiding argumentation. It might be objected that this does not actually constitute an evidential relation: the target of justification (y) in E_4 are decision options, i.e., they are prescriptive. If we thought that evidence always is supposed to be truth-indicative, and we do not believe that prescriptive claims are truth-apt, we would have to deny that E_4 is a use of evidence. However, I believe that it is in accordance with language use to say that we have evidence which justifies us to do something; for example, it seems perfectly fine to ask what evidence we have that might suggest that we should engage in mitigation. We might also express the difference between E_1 – E_3 and E_4 saying that while in the former, x is used to justify the belief that y is *true*, in the latter, x is used to justify the belief that y

is *right*.

This model is derived from examples related to policy-making; some of the cases we considered before do not neatly fit this picture. Recall the discussion of the introductory review from section 2.1: one of its functions was to highlight the importance of the novel research presented in the article by highlighting lacunae in the SoR. We could try and interpret this as the author presenting some facts about research results to justify belief in a hypothesis—that some research is missing—(E_2) and then proceeding with a normative argument for the decision option to engage in the research project. However, this seems to be an over-interpretation; some facts about the research are presented, and this is more or less directly used to justify the importance of the research in the article. Similarly for the fourth function mentioned there, i.e., signaling to reviewers the expertise of the author: while it would be possible to interpret this along the lines of providing secondary evidence (E_3) for the claims made in the article, it rather seems to be about the quality of the article, connected to the decision of publishing or not publishing it. It would, therefore, be better to provide an entirely different account for relevance in connection to reconstructions of the SoR that are not supposed to, in the end, provide evidence for some scientific claim.

Also, even remaining in the area of SoRs informing decision-making, not all types of facts can be used as evidence. Consider facts about the theories accepted by the scientific community, the methods employed in evidence collection, experimental setups, scientific modeling, statistical analysis, the values embodied in research practices, and even the material objects used in research, such as machines and instruments. All these elements contribute to scientific inference, but, as long as we do not subscribe to an extremely broad notion of evidence as justification, they should not be considered evidence. Sometimes, perhaps, the fact that some researchers used, for example, a questionable statistical tool in justifying their results could be considered secondary evidence. Also, it is notoriously hard to sharply distinguish scientific claims that I have described as evidence, and the theories that contain, imply, or are supported by them. However, it is plausible that coherence with a well-accepted theory can lend additional credibility to a scientific claim. Models can depend on, incorporate, or be tested against evidence, but sometimes they might not even be supposed to be factual or be believed in, but constitute rough heuristic tools paying respect to the contingencies of research—and still, they might be used to justify specific courses of action. Methods might feature in scientific inferences not as evidence, but as rules for how evidence can or should be used. For all of these elements, calling them "evidence" seems inappropriate. In this chapter, I can only hint at this complexity; analyzing what plays a role in scientific inference is a huge part of epistemology in philosophy of science. For the project presented in this book, it will have to suffice to remember that when we consider the (im-)balance of the SoR, this might not only be about the evidence, but also about these other aspects. In subsection 6.2.1, for example, with hermeneutical injustice I mention a concept which does not depend on an imbalance in the available evidence, but in the conceptual resources provided by the SoR.

Now that I have explored some of the central aspects of the SoR by distinguishing different uses of evidence, I will apply this model to the question of relevance, that is: when does a fact belong to the topic of a SoR?

2.3 Relevance

In the previous section, I have argued that, whenever someone wants to provide an account of the SoR on any topic, they will have to make decisions about which facts are important enough to be reported. Therefore, accounts of the SoR depend on normative considerations. The general, wide concept of SoR defined at the beginning of this chapter, however, does not—because it does not imply a need for selection. I have already said that the concept of SoR used throughout the dissertation has to be more specific than this broad idea of SoR because I am concerned with the SoR as a source of information concerning specific decisions and questions. Building on the initial definition, we could define this concept of a specific SoR as follows:

> **Def.: A state of research (SoR)** is the set of all facts about scientific research at a given time that are relevant for a specific topic.

Does this definition refer to a still infinite subset of facts about a particular facet of the overall SoR, similar to how the latter is just a facet of the state of the world? Is this concept of SoR in itself normatively laden, or are normative assumptions only introduced when trying to reconstruct them, i.e., do they arise from the contingencies of providing a written or spoken report for a specific audience? The answer to this question depends on how we understand the notion of relevance. In the following, I will argue that to determine if a fact about the research belongs to any particular topic, we already have to commit to normative assumptions, and thus the concept of SoRs used within this dissertation is dependent on a particular normative perspective on the topic at hand.

But what, in the first place, is a topic? To somewhat narrow this down, and because I am mainly interested in the effects of the SoR on decision-making, I will focus on topics as specific research questions or decisions to be informed. This understanding of the topic of SoRs, however, does not capture every colloquial use of the concept: after all, not all topics of SoRs are clearly connected to specific decisions. The SoR that is referred to in the IPCC's *Summary for Policy Makers* is explicitly concerned with policy decisions concerning climate change; the Cochrane Review aims at supporting health care decisions. But think of other types of reconstructions considered in section 2.1, such as textbooks for science students: clearly, the corresponding SoRs might influence all kinds of decisions about what lecturers might teach, how they present it, how students prepare for exams and more; but it is not clear if there is a primary one that could serve to give us a clear-cut idea of what research is relevant. Also, reconstructions of SoRs might have other effects on the audience besides influencing their decisions; it might instill a sense of wonder in them—think of science books for children—or simply satisfy their curiosity.

But how would we determine the scope of a SoR, given that we restrict ourselves to SoRs related to decision making? In the context of SoRs informing policy-advice, decisions occur at multiple levels. At the inner-scientific level, one has to consider which research hypotheses are to be accepted in light of the SoR. Let us return to the example of the systematic review given in section 2.1, or more specifically, to the method of meta-analysis. Meta-analysis can be considered a statistical tool for reconstructing particular aspects of the SoR on specific research hypotheses—in medical science, this will often be about the effectiveness of some medical intervention. Following standard interpretations of evidence-based medicine, not all evidence, but only evidence produced with methods above a certain quality threshold is supposed to be considered in meta-analysis. Authors like Stegenga (2011, 2018) have argued that this practice is actually a problem, as it violates the widely recognized principle of total evidence (PTE), which, very roughly put, says that when considering the degree of confirmation of a research hypothesis, we have to consider all available evidence. As the SoR is supposed to include every fact about the research about the topic, i.e., relevant to the decision, we can consider it to include also all available evidence, thus necessarily conforming to this idea of a PTE. But how does the PTE handle the question of relevance? Bengt Autzen (2016) reconstructs the principle in multiple steps, finally arriving at a preferred fourth version:

> Suppose data d1 are strictly logically stronger than data d2, then an inference about hypothesis H should be based on d1 if changing between d1 and d2 changes the evidential assessment. (ibid., 286)

In this formulation, what makes some fact relevant is its contribution to the "evidential assessment". This fits well with the aspects of SoRs discussed earlier: We have already considered three kinds of evidence which are used to justify scientific claims in section 2.2; also, I have argued that there is a range of other non-evidential facts about research which can feature in arguing for or against some claim, including but not limited to facts about theories, models and methods. All facts which can be used in scientific inferences (cf. Figure 2.2) resulting in relevant scientific claims should themselves be considered relevant.

But what in this criterion of relevance is normatively laden? I have been arguing that assessing the evidence is not just about applying some objective, mechanical rules to the data. As shown in Figure 2.2, at the very least, producing and interpreting the evidence also relies on, e.g., scientific theories. And as we can learn from Thomas Kuhn (1977), which *theories* are accepted—due to underdetermination—depends on the epistemic values of the research community, with legitimate differences in judgment concerning the individual members. A very obvious example is the theory of confirmation, which admittedly is a meta-scientific theory, but which is very much contested and directly influences the evidential assessment and thus the relevance of facts for the SoR. Secondly, the (non-)acceptance of scientific claims—or

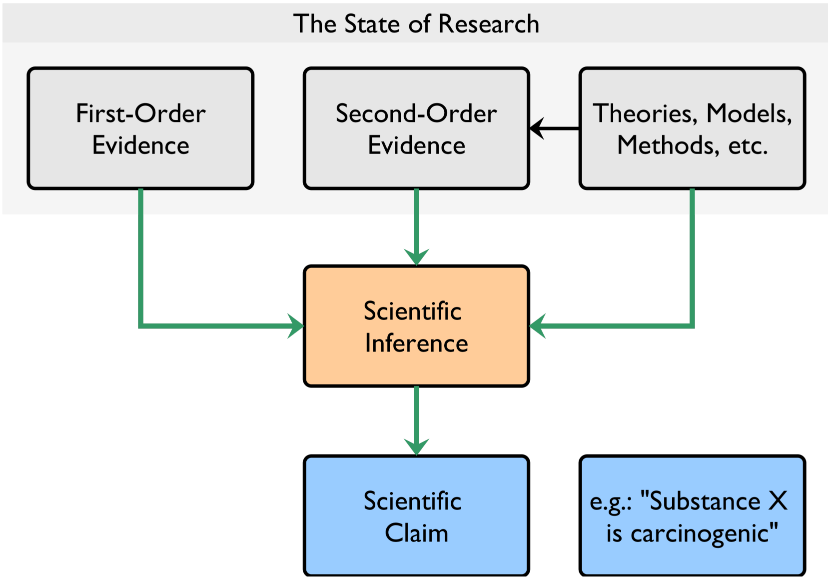

Figure 2.2: Relevance in Scientific Inference

the corresponding research hypotheses—is also a central point of discussion in the philosophical "values in science"-debate: the inductive risk argument as presented by Richard Rudner (1953) and also later by Heather Douglas (2000) deals with the role of non-epistemic values in accepting hypotheses. Some philosophers—like Gregor Betz (2013)—do not accept this argument and think that scientists should not make non-epistemic value-judgments in the justification of hypotheses. However, this stance itself is prescriptive on a meta-level, as it changes what should be considered as the basis of evidential assessment. Therefore, the explanation of relevance in terms of two levels proposed in this section explicitly is *not* a strict distinction between purely epistemic inferences on one and practical or ethical inferences on the other side.

But on a second level, the role of not simply normative, but genuinely ethical considerations on the scope of the SoR becomes even clearer: when determining what research is connected to a SoR with practical relevance, we also have to consider policy and other forms of decision making that can be informed by scientific research. This comes down to the question of how to determine which scientific claims are relevant to the SoR on some topic. Douglas (2012) claims that issues "such as whether a substance should be declared a 'known human carcinogen[...]'" (ibid., 144) are still to be considered problems of weighing the evidence, i.e., of scientific inference. However, the mere conclusion that some drug x causes cancer in humans—perhaps with some level of confidence or probability attached—is different from deciding to label it a carcinogen. The thing to be confirmed by the evidence, in this case, is not simply a scientific claim, but a question of policy with legal, financial, and other social consequences. Taking another example but staying in the realm of medical research, when a drug is considered for approval, not only should it be more effective than existing pharmaceutical interventions, the effectiveness also has to be weighed against the side-effects or harms the substance might produce. Only considering one drug, this can include a wide variety of effects, both positive and negative, which all come with their respective research hypothesis. Still, the question of what should influence the decision and what not appears relatively clear: if there is evidence for a positive effect, this speaks for the approval of the drug; if there is evidence of harmful effects, that speaks against it. However, what is going on here is not scientific inference, but practical argumentation.

Consider the much more complicated issue of the public debate about climate change policy—that is, the question of how and with what concrete measures to react to the rising temperatures resulting from anthropogenic emissions. Of course, climate change skeptics even have questioned if there is any such thing as anthropogenic climate change. The answer to this question could still be put as a single claim in line with the first level of decisions discussed above. If, however, we are to decide which action to take in the face of it, this implies a very complicated decision problem that could be informed by scientific results at many different points: there is research on all kinds of different aspects of the climate system and its response to changes both on global average and in local settings, there is the question of what these changes in the climate system might cause, including feedback loops, effects on agriculture, economic costs, or loss of life due to rising sea levels, food shortages, and social conflicts. Furthermore, the use of certain interventions such as geo-engineering presupposes the research and development of said technologies. In this way, even the possibility of some decision options depends on previous research. Lastly, also questions of technology assessment of interventions feature in the overall decision situation.

But how are these arguments connected to the relevance of scientific claims? At least some of the arguments will include empirical premises, which could be supported by the results contained within an SoR. However, as these have to be action-guiding arguments, they will also have to include normative bridge-principles, which allow us to make inferences to prescriptive conclusions from descriptive research hypotheses, such as in the simple scheme shown in Figure 2.3.

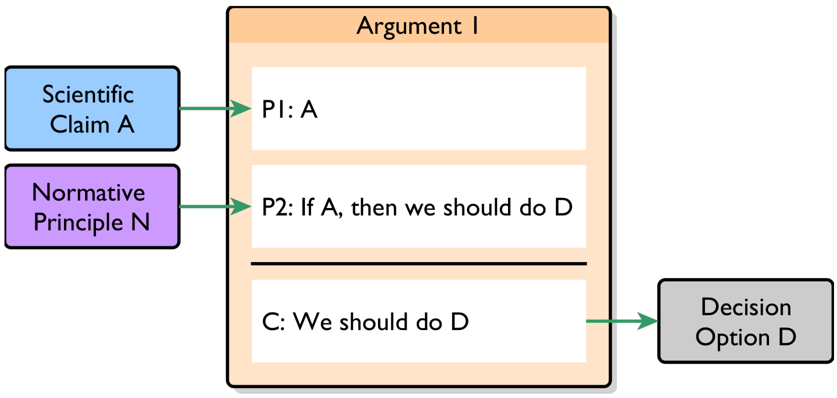

Figure 2.3: Relevance and Bridge-Principles

Research hypotheses, and with them the facts that are used to justify or reject them, are only relevant to a decision if we also accept these normative principles. In the climate change example, research on, e.g., the effects of stratospheric aerosol injection (SAI) on the beauty of the earth's skies—SAI being a proposed climate-engineering technique that might cause a whitening of the earth's sometimes blue skies (Robock 2008, 16)—is only relevant if we consider the loss or change of such beauty to be relevant to the decision about climate response policies. Therefore, when considering the SoR as it relates to a decision problem to be informed by scientific research, we must add another inferential level downstream from the scientific inferences (cf. Figure 2.4).

Relevance on the second level is determined by the connection of scientific claims to the decision option, which is established by action-guiding arguments. In the example of toxicity assessment used by Douglas, research evidence which is used to justify the finding that some substance x is carcinogenic becomes relevant because this

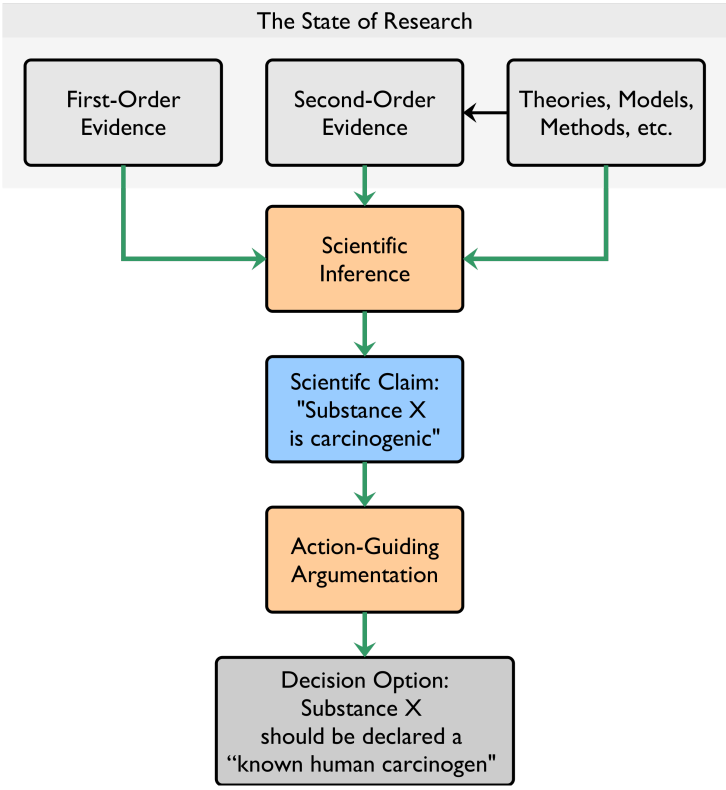

Figure 2.4: Two levels of inferences

claim may be used in a premise in an argument for officially labeling the substance a “known human carcinogen” (Douglas 2012, 144).

Just as with a multitude of scientific evidence supporting a claim, various arguments may support or attack a specific course of action. Figure 2.5 schematically shows a more complex argumentative network: in this model, research hypotheses support (green arrows) or attack (red arrows) arguments, which can be interpreted as the scientific claim being identical—in the case of support—or contradicting—in the case of an attack—a premise in the argument it is connected to. With arguments supporting (or attacking) an element, their conclusion is to be thought of as being identical (or in contradiction to) a subsequent premise or, in the end, the decision option under consideration. Note some other interesting features shown in the model: firstly, research hypotheses do not have to support or attack the action-guiding arguments directly, but there might be intermediary steps. For example, recall the example

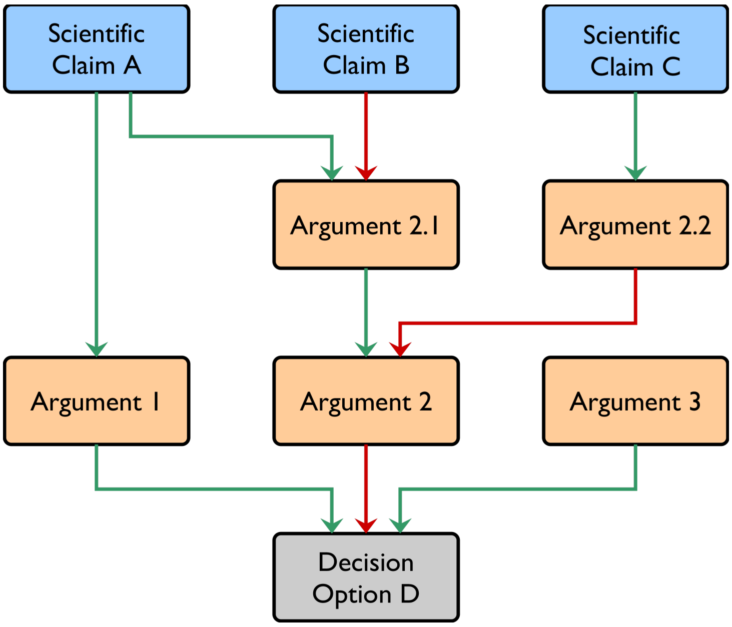

Figure 2.5: Scope and Networks of Argumentation

of paleoclimatic records as evidence for the claim that changes in atmospheric CO_2 concentration play an important role in glacial–interglacial cycles (cf. Scientific Claim A in Figure 2.5). There, this argument is not clearly directly connected to an argument suggesting a climate response policy, but perhaps it first might be used to argue that human emissions of greenhouse gases exacerbate global warming (cf. Argument 2.1), this conclusion supporting another argument (cf. Argument 2) which, for example because of the resulting rise in sea levels threatening island nations, concludes that we should not engage in a "business as usual"-policy, continuing with CO_2 emissions (cf. Decision Option D). Secondly, the model shows that an accepted scientific claim can both be used as evidence for and against a decision option at the same time: in our example, the claim about the role of CO_2 (cf. Scientific Claim A) might support another argument (cf. Argument 1) which argues that we should continue with business as usual (cf. Decision Option D), because, for example, climate change might benefit agriculture in Greenland or Canada.

This structure also emphasizes that when deciding how to act, we must weigh the different arguments against each other. In drug approval, for example, we might have to weigh arguments for approving the drug—e.g., concluding in the safety of the drug—against arguments against approving the drug—e.g., concluding in it not being safe (cf. Figure 2.6). Relevance, then, runs in the opposite direction of the inferences (cf.

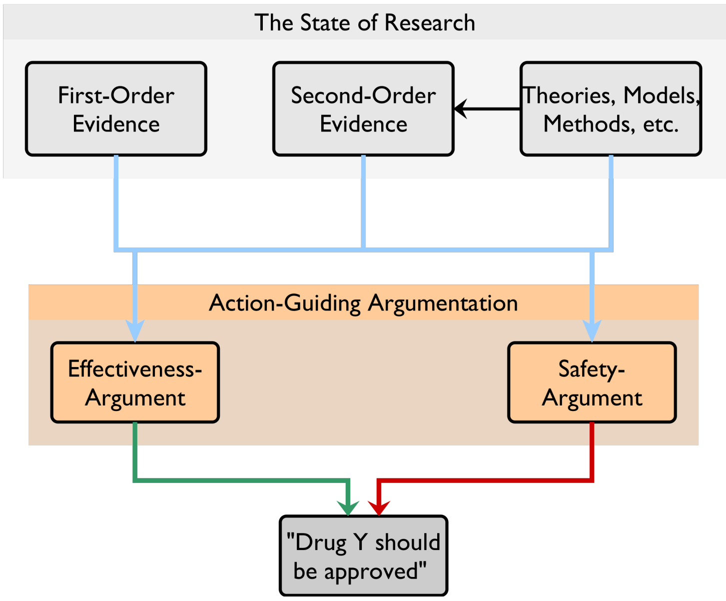

Figure 2.6: Weighing Arguments

Figure 2.7): Starting with the decision option, we can point out relevant arguments in so far as they support *or* attack the decision option, and relevant scientific claims in that they support or attack relevant arguments[7]. Also, the more weight we assign to a specific argument, the more relevant becomes the scientific evidence that can support or contradict the hypotheses used in their premises.

Relevance also depends on the decision principles that are used by decision-makers or accepted by the participants in the discussion. Normative arguments such as the ones envisioned above are often only pro tanto, i.e., it is not obvious how to aggregate the various arguments and actually arrive at a decision. Depending on how we envision this step, the relevance of hypotheses will change: For example, in classical cost-benefit analysis (CBA), we can only use information about events to the

7. This depiction of an inferential network as a directed graph is not supposed to imply that the scientific process works in a linear fashion: decision problems may inspire new research just as new research may lead us to consider new practical or ethical problems, for example.

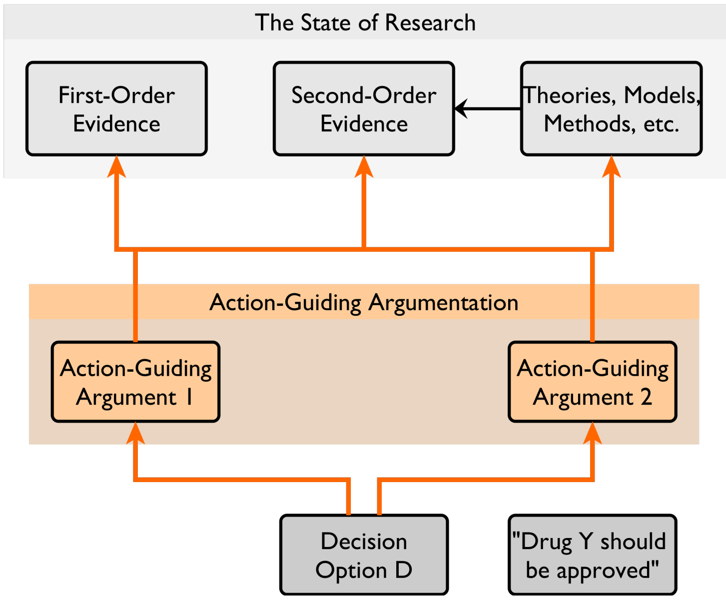

Figure 2.7: Inferential Network and Relevance

occurrence of which we can assign some numerical probability. Otherwise, we will not get any expected utilities. If we, however, favor a form of precautionary decision principle such as maximin, scientific claims that deal with uncertain events might become relevant; however, only those claims which are about the worst consequences of each decision-option are of relevance.[8]

One might object that the dependence of relevance on normative positions is an artifact of idiosyncratic views embedded in reconstructions and that the SoR itself should not depend on any particular ethical stance but take into account all the arguments that apply to the debate in question. This, however, results in a chain of questions such as "Is it supposed to be all arguments that actually have been put forward at a specific point, or all arguments that could be advanced by the involved parties?" or, respectively "If it is supposed to be all the arguments that involved parties could have provided, is this about the parties actually involved or about all

8. Of course, both in CBA as well as concerning precautionary principles, there are many alternative proposals; for example, the use of maximin might be though to require that the envisioned benefits are relatively small compared with the possible costs, broadening the scope of the SoR. However, this multiplicity of approaches in decision theory only emphasizes the effect of decision-principles in determining relevance.

the parties which would be affected by the decision in question?" which would again require substantial normative argumentation. Even if we could find a way to avoid the dependence of the scope on any particular set of normative premises, this would only make the SoR neutral towards the possible positions, and not ensure its independence from normative argumentation altogether.[9]

Other critics might ask why one might need this explanation of relevance at all. After all, I mentioned hat Kitcher's notion of significant truth already distinguishes truths simpliciter from interesting truths. His ideal of well-ordered science (cf. Kitcher 2001, chapter 10) can be understood as an attempt to organize research in a way so that it produces interesting, that is, significant results. The graphs that model relevance in this chapter are reminiscent of Kitcher's significance graphs. These are graphical representations of the relative significance of research items, relative, that is, to some higher-order questions. Figure 2.8 shows Kitcher's significance graph for the question "How do organisms develop?". Kitcher explains the idea behind these significance graphs as follows:

> Fields of science are associated with structures I shall call *significance graphs* that embody the ways in which their constituent research projects obtain significance. A significance graph is constructed by drawing a directed graph with arrows linking expressions, some of which formulate questions that workers in the field address, others encapsulate the claims they make, yet others that refer to pieces of equipment, techniques, or parts of the natural world (figures 1, 2). The significance graph reveals how to explain the significance of various items—where "item" is an all-purpose term for questions, answers hypotheses, apparatus, methods and so forth. One would account for the significance of the item to which the arrow points on terms of the significance of the item to which the arrow points in terms of the significance of the item from which it comes. Arrows thus display the inheritance of scientific significance. (ibid., 78)

9. However, there is a more serious argument to the effect that not all SoR might be dependent on normative questions in the same way: I have only focused on SoRs where the topic is about informing decision-making. By nature, this requires answers to normative questions, and it does perhaps not come as a surprise that the scope in these cases is connected to normative argumentation. There might be SoRs, such as reviews giving an overview over all the research in a specific field of theoretical physics, without any concrete applications besides satisfying epistemic interests, where a connection to action-guiding arguments is not apparent.

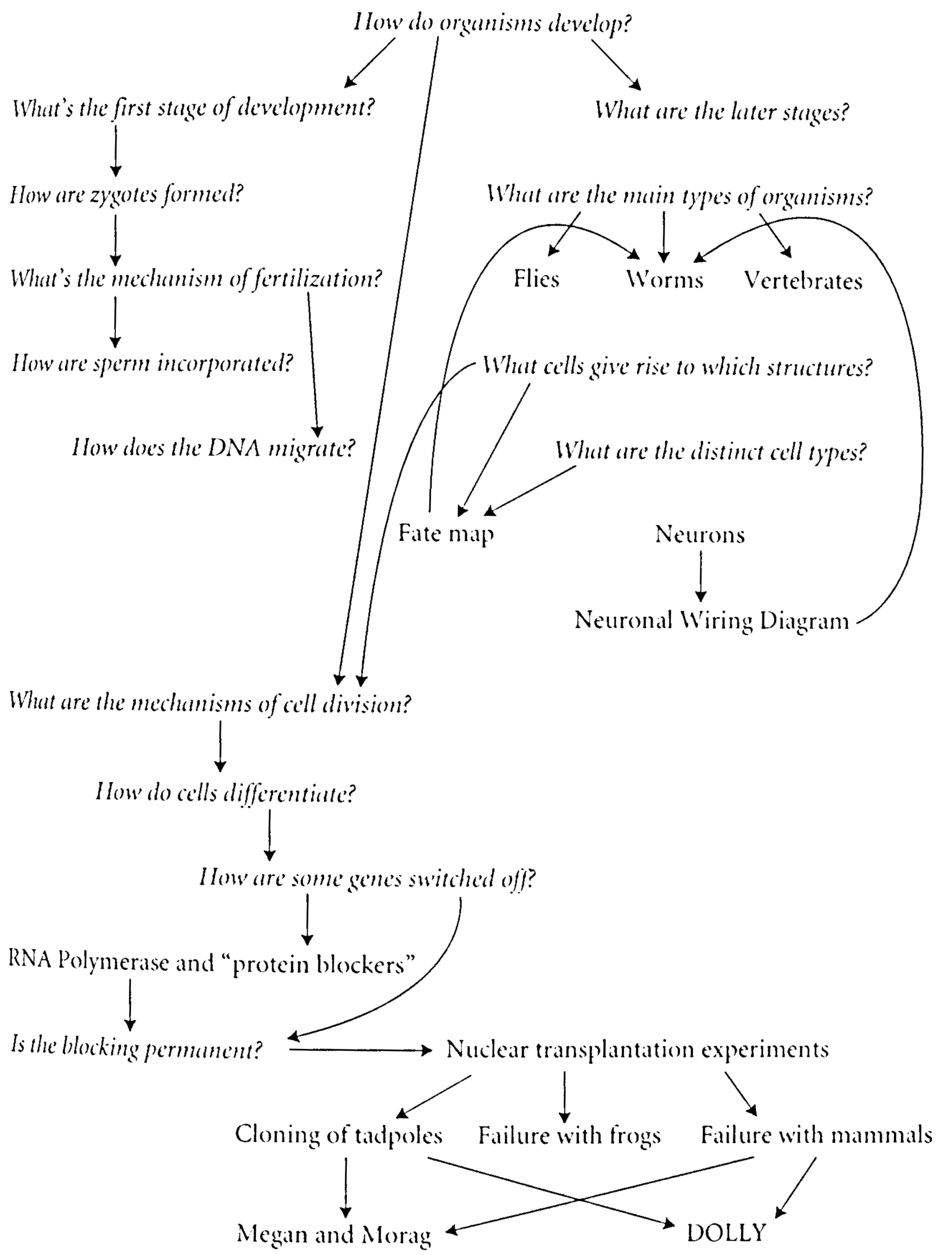

Figure 2.8: "An extremely partial depiction of the significance graph for developmental biology", reprinted from Kitcher (2001, 79).

Obviously, there are some similarities between Kitcher's approach and the one in this chapter: both are about the relative importance of products of research, ordered hierarchically towards a main object of interest. Also, both relevance in this chapter and significance in Kitcher's work are explicitly relative to normative considerations: "Like maps, scientific theories—or better, significance graphs—reflect the concerns of the age. There is no ideal atlas, no compendium of laws or 'objective explanations' at which inquiry aims." (Kitcher 2001, 82)

So why prefer the notion of relevance over that of Kitcher's significance? There are some important differences between Kitcher's account of significance and relevance to a SoR: firstly, significance implies that the end-goals of research are actually ethically desirable. "Think of a problem for investigation as arising when some entity of a specified type is sought. Problems worth pursuing can be labeled as **significant** [my emphasis]" (Kitcher 2011, 105). That is, the significance of the end-goals to which the research contributes is an intrinsic ethical value; relevance, on the other hand, is just about the importance of some item of research for some goal that someone might have, depending on their normative principles. This difference will be of importance in the second part of the dissertation.

Secondly, what exactly the relations between the nodes in Kitcher's significance graphs entail is left rather vague. Obviously, an item that is significant for something else has to be somehow important for it, but how is this importance determined? Kitcher provides one explanation in terms of providing solutions to the significant problems:

> Those problems are *adequately solved* when an item is produced that is close enough to the type sought to serve the purposes that confer significance to the problem. If the problem is to answer a question, an adequate solution is a statement "true enough" to enable those who have it to achieve whatever ends made the question significant. If the problem is to produce a new vaccine, an adequate solution is one providing acceptable protection against the pertinent disease. If the problem is to develop a new technique, an adequate solution is one allowing people to proceed sufficiently successfully in the contexts of intended use. (ibid., 105)

However, that is obviously not all there is to be said about the significance relation: in the significance graphs, multiple arrows may point towards the same item, and this does not mean that at each of the origins you will find a solution; rather, it seems that, taken together, the items at the origin of the arrows contribute to that solution. To be more explicit of how we can think of the relation between items, I have chosen an inferential model, which conceives of relevance in terms of support and attack relations between premises and conclusions of arguments. This also allows me to clarify how normative considerations influence the relations: it flows from both the role of normative statements as bridge-principles in these arguments and the rules of inference themselves. This is not to say that Kitcher's more open concept does not also have its advantages: the argumentative framework used in this dissertation leans towards a rational reconstruction of the scientific process and its relation to decision problems; Kitcher very rightly emphasizes that not only cognitive achievements in terms of scientific findings, but also material aspects of science such as techniques or physical products—be it vaccines or instruments—play an important role. This is a compromise: As my focus is on the question of how scientific information can inform decision-making, I focus on the cognitive outputs of science. Other important aspects of research, such as the development of drugs, are indirectly accommodated as answers to questions such as "How should we treat disease *y*?", or in the context of scientific inferences, as part of the non-evidential contributions to the process (cf. Figure 2.2).

In conclusion, although perhaps to varying degrees, the scope of both specific SoRs in themselves and their reconstructions depends on normative attitudes, be they about epistemic values in the acceptance of hypotheses or moral principles and decision theory in policy-relevant decision-making. What makes facts relevant to an SoR informing decision-making is their usefulness in either the action-guiding argumentation about the decision options, or the underlying scientific inferences justifying the relevant research hypotheses.

In this chapter, I have introduced the concept of SoRs and their reconstructions and pointed out what kinds of facts they include and how to determine which of them are relevant to decision-making. In the next chapter, I will discuss cases in which the available facts thus determined seem to suggest a problematic imbalance in the SoR.

Chapter 3

What is Imbalance?

In the previous chapter, I have defined the concept of "State of Research" as the totality of facts about the research on a certain topic. I have also—by way of the discussion of relevance—highlighted a connection between such SoRs and action-guiding arguments in the context of decision-making. The main goal of this dissertation was to give a philosophical account of problems that can occur at this intersection between the accumulated findings of science and the subsequent decisions which rely on the former for evidence. My focus, in this chapter, is on phenomena that go beyond problems with any particular finding: Individual results being unreliable, inadequately justified or flat-out erroneous are of course problems for decision-makers. But while questions about the justification and reliability of research findings have been widely discussed within philosophy of science, the perspective on the aggregated information provided by science paves the way for a discussion of further, novel problems. This requires a different conceptual framework, which can explain why some SoR might be problematic, even if the individual pieces of research are not criticized as such.

In his 2007 book, Roger A. Pielke Jr. introduces the different roles of scientists can take up in policy-making with the metaphor of a local recommending a dinner place to an acquaintance visiting town (cf. Pielke 2007, 1). Imagine, instead, a quite similar, but somewhat different scenario:

Instead of one person recommending a restaurant to another, the visitor—standing in for the policy makers—has no local friend to rely on, but luckily she is in the possession of the modern wonder we call a smartphone. She opens up her favorite application for urban navigation, and types in "dinner". Our hungry traveler will probably be rewarded by some symbols popping up on a digital map of the city, indicating places where one might enjoy an evening meal, complete with the opinions of previous customers, a rating—perhaps in the form of 0–5 stars—and a number indicating how many people have reviewed this locality. What problems could occur with the information provided here? Firstly, we might again be concerned with the reliability of the individual reviews given: we might wonder if the ratings actually reflect the quality of the various restaurants, or rather, the visitor's preferences concerning the dinner. There could be all kinds of issues with the ratings and comments: reviewers might not be competent—perhaps they have very limited experience with food, service or other relevant aspects of restaurant visits—their interest in dining places might be very different from our visitor's, or they might favor specific places for reasons that are not about quality at all—maybe some reviewers have personal, positive or negative connections to some of the restaurants' owners, have been paid to give good or bad reviews, or perhaps the owners themselves use the app to advertise their establishments fraudulently.

But the individual content of the ratings and comments are not the only information provided in the example: the visitor can also see how many ratings there are for each place. This she might take for secondary evidence about the reliability of the ratings, meaning that for two places with four stars, one having hundreds, the other only a handful of reviews, the former will appear as the safer option. Other places might be new, and therefore not have any reviews yet, while some other restaurants that the visitor might see in the streets might never appear in the app, perhaps because the owners do not have the technical know-how to register it there. That is, we might not only doubt the reliability of the restaurant guide because of the quality of the reviews, but also because of their quantity. Here, however, the importance of quantity is not just about the simple fact that the app would be more useful if there would be more reviews in total, but also about the question of how they are distributed among the individual options: it is not just that more reviews for one place would make the information more accurate, but that it becomes relatively less attractive compared to other restaurants that have attracted many reviews. This latter problem is especially evident when we suspect some systematic influence behind the distribution: perhaps some part of town has worse wireless internet than the rest, and although that area is very close to the visitor and full of excellent places, people can't access the app there, leading to a lack of reviews. Maybe the algorithms behind the app itself give priority to certain places, perhaps by actively asking the user to leave comments for certain establishments, but not for others.

Obviously, this toy example does not map directly onto the connection between science and policy-making; science, after all, does not give out ratings concerning policy options on a five-point scale. However, what we can take away from this analogy is the idea that not only the quality but also the quantity and distribution of research can matter when discussing the reliability of the SoR in some area.

One related concept, which we already encountered in the discussion of reconstructions of the SoR in section 2.1, is the idea of lacunae; that is, some piece of information which we deem important might be missing—which does not imply that any of the results we already have are flawed. Like with the example of the restaurant guide, not the quality of individual results, but their distribution is in question. Note that the latter can influence the former: for example, consider the concept sometimes referred to as "value-laden blindspots" (cf. Elliott and McKaughan 2009; Bueter 2015), where lacunae in the research landscape negatively affect the epistemic accuracy of decisions concerning theory or hypothesis acceptance. This is also generally true for research synthesis in secondary research, in which imbalances in the SoR would also distort the result of the synthesis as an individual piece of research. Thus, the idea that such problems with distribution can, in principle, occur independently of problems with individual results, should not be taken to imply that this always has to be so, or is especially common. Both issues might often work together to further aggravate problems in a particular case.

In the following, I will collectively refer to problematic phenomena that are connected to the distribution of research as **imbalances in the SoR**. While the term "lacuna" implies that there is some gap, i.e., there is no research of a particular kind, the term "imbalance" is more broadly construed, in that it generally stands for cases where there is a disproportionate amount of research of some type(s)—"disproportionate" allowing for both the meaning "too few" as well as "too much". In section 3.1, I describe central examples of phenomena which gave rise to the idea of imbalances in the SoR and discuss similarities and differences between them. In section 3.2, I provide a theoretical approach to the concept of imbalance, and discuss its relation to other terms connected with the examples presented, such as independence and objectivity.

3.1 Imbalance of the SoR in the Literature

In this section, I will provide an overview of examples for imbalances in the SoR taken from the meta-scientific literature, both in terms of concrete real-world cases and more abstract descriptions of mechanisms that might lead to imbalance. This list provides the bulk of the material that will be analyzed in the rest of this dissertation[1]. It follows the lines of various categories that may be used to conceptualize imbalance. Below, I will first guide through the examples grouped by what type of distribution is considered the problem—lack of research or its overabundance and a variety of sub-types that I will introduce along the way. After that, I will discuss various other ways to categorize different kinds of imbalances.

3.1.1 Lacunae in the SoR

The idea of lacuna discussed above defines the first type of example for imbalance in the SoR. There are many structural features of scientific research that have been criticized for causing a lack of particular kinds of research. I here distinguish three ways in which research might be lacking: firstly, research of a particular kind might never have been done in the first place—**undone science**—, secondly, research might be available in principle, but is ignored because it is not considered relevant—what could be called **excluded research**—, and thirdly, research might have been done, but is not made accessible to the parties relying on the SoR—**inaccessible research**.

Undone science. The term "undone science" was introduced by sociologists Frickel et al. (2010) to describe "areas of research that are left unfunded, incomplete,

1. See Table 3.1, p.86, for a full list of the major examples.

or generally ignored but that social movements or civil society organizations often identify as worthy of more research" (ibid., 444). Frickel et al. (2010) themselves give multiple examples for undone science, the first of which comes from the context of environmental science: With the "chlorine sunset controversy" (ibid., 449)—a conflict about the manufacture of chlorine compounds in the Great Lakes region—the authors discuss the idea that **regulatory paradigms** can lead to undone science. The conflict is traced back to competing policy frameworks for toxicology and the consequences of these regulations for the research that has been done. In both the USA and Canada, the dominant regulatory regime follows the idea that the burden of proof concerning the toxicity of substances lies with the public—a position the authors identify as the "risk"-paradigm—which is challenged by critics who demand that industry should have to provide evidence that the compounds they produce are safe—which is dubbed the "precaution"-paradigm. The dominant risk-paradigm, or so the authors argue, has led to a SoR where government research focuses on the ad-hoc identification of unsafe chlorine compounds while industry research which aims at developing new chlorine-based chemicals to replace substances which are found to be problematic. This is criticized because, due to resource constraints, government research can never check all the compounds that are in use—it is, as the authors say, "undoable science" (ibid., 453). Also, research projects which would investigate the dangers associated with chlorinated chemicals as a class or aim at developing alternatives to a chlorine based industry altogether have been neglected. This is seen as especially problematic because these alternative lines of inquiry, while they might put more pressure on the industry, may also contribute more to the protection of consumers and the environment.

In philosophy of science, different phenomena constituting undone science have been discussed in connection with medical research: James Robert Brown (2008) discusses the idea of "**unpatentable research**" (ibid., 197). He argues that the current funding structure of medical science leads to "skewed research aimed toward patentable solutions, away from exercise, diet, environment, and so on" (Brown 2016, 525). Because of the costs involved with medical trials, research relies heavily on the support of the pharmaceutical industry. And, or so the argument, financially interested sponsors will always support research projects which conclude in patentable innovations such as new drugs. These will thus receive more attention than alternatives such as changing diet, although the latter might potentially be at least as useful in promoting health—but which cannot be monetized to the same degree. Similarly, Reiss and Kitcher (2009) describe the phenomenon of **neglected diseases** (cf. also Flory and Kitcher (2004)). Here, the imbalance is not about which types of intervention can be patented, but about which diseases are the object of research in the first place. The idea is that diseases mostly affecting people which are not able to pay for expensive treatments might receive less attention than other afflictions. These issues are structural in nature, as they arise from the way in which the interactions of science, government and the private sector are regulated: it is not just about some individual decision of what should be researched, but about how funding for research can be acquired and who can come up with the money for large scale clinical trials, about patent laws, and other general features of the science system.

Undone science is also described at more basic levels of biomedical research. Consider the issue of **ignored genes**: researchers like Edwards et al. (2011) have claimed that genetics research has had an unduly limited focus. While a large proportion of research has investigated only the same "50 proteins that were the 'hottest' in the early 1990s" (ibid., 164), "the human genome encodes more than 500 protein kinases, of which hundreds have been shown to have genetic links with human disease" (ibid.), most of which apparently have remained ignored by scientists. The authors give many reasons for why that might be the case:

> [S]cientists are wont to "fondle their problems": they have a natural tendency to dig deeper into their areas of expertise. Plus, funding and peer-review systems are risk-averse; funders and reviewers alike are less willing to support research on unstudied proteins, for which it is often harder to explain the rationale and significance. Moreover, the time frames associated with academic promotion and training encourage researchers to focus on systems that are likely to generate results rapidly, and for which research infrastructure and methods are already available. (ibid.)

This list of reasons is especially interesting because it shows that not only outside influences like the interests of sponsors or policy can lead to undone science, but also the demands of academic careers and thus the interests of the scientists themselves. It should be highlighted that there is an element of path-dependency here as well: for example, the authors emphasize that being able to do previously undone science often hinges on the prior development of research tools that allow the study of, for instance, a new protein. Thus, once tools for a specific area of research have been developed, it is very costly and, therefore, also difficult to change the course of the SoR.

Excluded research. Not only the context of research, but even the very standards of scientific communities can be a source of lacunae in the SoR. Consider the "Evidence-Based Medicine" (EBM)-movement mentioned in section 2.1: Institutions such as Cochrane endorse a focus on systematic reviews & randomized controlled trials as the most valuable form of evidence in medicine. There are good reasons for placing great value on meta-analysis and RCTs in medicine—because these research designs can help avoid certain biases associated with other ways of collecting evidence such as observational studies or, e.g., narrative literature reviews. If, however, the **evidential standards of EBM** lead to an exclusive focus on what are considered the best kinds of evidence, other types of evidence will become irrelevant and thus no longer count as part of the SoR. As I have argued in section 2.3, whether some fact about the research is to be included in the SoR or not depends on its relevance for decision-making. In the systematic reviews—which are supposed to provide the primary source of evidence for decision-makers in health-related questions (cf. Higgins and Green 2008, 6)—, almost exclusively randomized trials are supposed to be considered (ibid., 90). Other research, for example, based on individual patient histories, would have no bearing on decision-making, and would, therefore, have to be considered irrelevant. However, authors like Nancy Cartwright (2012) and Jacob Stegenga (2018) have argued that RCTs and meta-analysis alone are not enough to form a reliable basis for policy and individual-decision-making, and we should allow for greater evidential diversity. These other, neglected types of research could become undone science because researchers and funding agencies might be reluctant to invest in projects whose results, in the end, will remain ignored. While this might be a worry for the future, at the moment is a large body of data and scientific papers that use and report other types of evidence. When the standards of EBM are criticized because these sources of evidence are not considered part of the SoR although they might contribute to decision-making, they constitute what I have called excluded research above.

Frickel et al. (2010) include a second example besides the aforementioned regulatory standards: the case of **air-monitoring-standards**. While this is also presented as an example for undone science, I argue that it should rather be considered excluded research:

> In communities adjacent to refineries, power plants, and other hazardous facilities, known as "fenceline communities", residents suspect that facilities' emissions of toxic chemicals cause serious illnesses. However, there is a dearth of scientific research that could illuminate, in ways credible to residents, the effects of industrial emissions on community health [...]. (ibid., 454)

While this void in the SoR could, in principle, be filled by the involvement of activists who collect air pollution data locally using so-called "buckets", the relevance of these contribution hinges, once again, on the regulatory, i.e., action-guiding framework: "[A]mbient air standards are typically expressed as averages over a period of hours, days, or years. [...] Bucket data, in contrast, characterizes average chemical concentrations over a period of minutes." (ibid., 455) Therefore, the parties involved—such as industrial companies and government agencies—may dismiss the activists' efforts as irrelevant. Only if the regulatory standards for air pollution that determine what evidence can have an impact on policy-making allow the inclusion of such contributions from citizen science, will they become relevant, and thus part of the SoR.

Inaccessible research. Besides excluded research, there is another type of lacuna that also refers to research which has been done, but does not form part of the SoR: Research that actually *was* done may remain unpublished, only published with a time-lag (Song et al. 2010, 24-26) or perhaps only available from the so-called "Grey literature" (ibid., 26-29) where it remains ignored[2]; it is what Chan et al. (2014) call "inaccessible research". The perhaps best-known phenomenon of this type is **publication bias** as first discussed by Theodore Sterling (1959) and which is sometimes

2. While these two latter forms of dissemination bias do exist both in analogy to publication bias and suppression, for the sake of brevity, I here focus on cases where some finding is actually not made publicly available at all and is not just published with delay or in media of questionable scientific repute.

also called "the file-drawer problem" (Rosenthal 1979): "[I]n fields where statistical tests of significance are commonly used, research which yields non-significant results is not published" (Sterling 1959, 30)—or at least less likely to be. Beginning with Sterling's article, there have been many empirical studies confirming the existence of publication bias in the health sciences (Song et al. 2000, 7-21; Song et al. 2010, 9-19). The sources of publication bias are, as the following definition suggests, manifold:

> Publication bias is specifically defined as "the tendency on the parts of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or strength of the study findings". (Song et al. 2010, 2)

That is, while publication bias is always about some type of result being less likely to be published, it is not specified if this occurs because that kind of research is not *publishable*—i.e., reviewers or editors are unlikely to accept them—or because nobody even attempts to publish them—meaning that the authors themselves do not even submit their findings. Song et al. (2010) find that, more often than not, actually the latter kind of decision is the source of publication bias (ibid., 41-50). This finding, however, might also be attributed to a feedback effect, where, once scientists have learned that certain kinds of results will most likely not be accepted by editors or go through the review process, they decide not to waste their time submitting such findings in a form of anticipatory obedience. Or, perhaps, they might not even pursue projects that are unlikely to produce publishable results, which would create yet another form of undone science as well. As Nobel Prize winner Randy Schekman writes in an opinion piece published in *The Guardian*:

> A paper can become highly cited because it is good science – or because it is eye-catching, provocative or wrong. Luxury-journal editors know this, so they accept papers that will make waves because they explore sexy subjects or make challenging claims. This influences the science that scientists do. It builds bubbles in fashionable fields where researchers can make the bold claims these journals want, while discouraging other important work, such as replication studies. (Schekman 2013)

While the original concept of publication bias has focused on non-significant vs. significant results, in the newer literature on publication bias, the categories of what is thought of as neglected versus favored have broadened considerably:

> Study findings are commonly classified as being statistically significant or non-significant. In addition, study results may be classified as being positive or negative, supportive or unsupportive, favoured or disliked, striking or unimportant. (Song et al. 2010, 2)

While I do not want to suggest to restrict the scope of the notion of publication bias to the question of significance alone, I think it is helpful to distinguish cases where the divide is in the quality of the result itself—significant versus non-significant, but also, e.g., confirmatory versus non-confirmatory—from cases that are about whose interests the results favor. This allows us to clearly distinguish the phenomenon of **suppression** of research findings from publication bias[3]. Suppression refers to cases where some party with vested interests—mostly sponsors with financial interests—works towards keeping some results from being made public. In philosophy, this idea often appears in connection to the more general debate about the commercialization of research and the role of research as intellectual property (cf. Brown 2008, 192-194; Carrier 2008, 219-221; Christian 2017). Suppression of findings can occur in terms of selective reporting of outcomes within a single study, but also to cases where whole studies go unreported.

3.1.2 Overabundance in the SoR

Lacunae are not the only type of imbalance that has been discussed in the meta-scientific literature: sometimes, it can also be problematic if there is too much research of a kind in the SoR. At first, this might seem perplexing: as long as the individual results themselves are unproblematic—the possibility of which I have made a central precondition of calling something an imbalance in the SoR—, how can having more research ever hurt? I do not refer here to just the other side of lacunae: the fact that, given limited resources, favoring some type of research always comes of the cost at

3. This will be of some importance in chapter 5, where the former is more directly connected to the idea of practical risk rather than alethic risk.

neglecting another. Instead, I want to point out three distinct types of imbalance where the prevalence of some kind of research itself is considered a problem: firstly, spending resources on some project that does not seem promising in terms of practical or even just purely epistemic benefits can be considered **research waste**. Secondly, a specific topic might be so much in the focus of research, that the importance of the subject under investigation is overestimated either by the public or even within science itself—**overemphasis**. Thirdly, although this is different from other forms of imbalance in important respects, I want to comment on the idea of **normatively inappropriate dissent**.

Research waste. Sergio Sismondo (2008) discusses several ways in which industry funding might negatively affect medical science. One of the causal structures he identifies is what he calls cases of "**Multiple trials with predictable outcomes**": he claims that drug trials funded by pharmaceutical companies might sometimes "be designed to test an already-studied drug in a way known to be effective, on a population for which it is known to be effective" (ibid., 3). While one could interpret such trials in terms of replication—making them useful in establishing the reliability of earlier findings—as the qualification of "predictable" research suggests, the financially interested sponsors will hardly be interested in undermining their own established products. Therefore, they will only sponsor such trials if they are quite sure that there will be no surprises. Such studies would then have to be considered research waste. Sismondo claims the reason for *why* sponsors might be interested in spending money on these trials is better understood as a form of marketing, "designed more to 'familiarize' physicians and patients of products than to produce novel knowledge" (ibid.). The idea is that, by producing research which is then disseminated to medical professionals, the salience of their products among physicians will increase. They, in consequence, might be more likely to recommend a drug they often read about to their patients. This example, therefore, also constitutes overabundance of the second type: it's a case of overemphasis.

Overemphasis. The term "overemphasis" is meant to denote cases where the focus on a particular topic is perceived as an indicator of the importance of said topic, while there is no actual justification for that belief. Recall the example of the restaurant guide from the beginning of this chapter: sometimes it is a valid heuristic to assume that a restaurant which has many reviews is a better bet than an alternative, which, ceteris paribus, has fewer reviews—perhaps because the food there is so good that it attracts many customers who then review the place. In other cases, where, for example, the reviews are sponsored by the owners, relying on this quantitative information might skew the perception of the app's users.

While waste research and overemphasis are a likely pairing—because the label of waste research already indicates that there is no good reason to focus on it—one does not always imply the other. Take the discussion about the alleged "**'Pause' in Global Warming**" (Lewandowsky, Risbey, and Oreskes 2016): the debate revolves around claims made by critics of mainstream climate science according to which the increase in global temperatures has stopped, lessened or at least not been as high as predicted by climate models. These claims made by dissenting individuals had a considerable impact on climate science, with a large number of articles having been published trying to explain the apparent hiatus in climate change. The problem—according to Lewandowsky, Risbey, and Oreskes (2016)—is not that this research is not in itself beneficial: "The body of work on fluctuations in warming rate has clearly contributed to our understanding of decadal variations in climate." (ibid., 729). The problem is that "by accepting the framing of a recent fluctuation as a pause or hiatus, that research has, ironically and unwittingly, entrenched the notion of a pause (with all the connotations of that term) in the literature as well as in the public's mind" (ibid., 729). The attention and responsiveness of the researchers to climate change critics can falsely be interpreted as an indicator that there are serious problems with the models of climate science that warrant skepticism about anthropogenic climate change. However, this attention—or so the authors—was not an effect of the seriousness of the problem, but of the scientists feeling that they need to respond to the widespread criticisms in the media and beyond:

> Here, we suggest that a contrarian meme can find entry into the scientific community simply by exploiting scientists' commitment to explanation and to responding to intellectual challenges. [...] In a world in which contrarian claims in the media and other public arenas are overrepresented [...], scientists

> may feel the need to respond to these claims. (ibid., 728)

Normatively inappropriate dissent. Dissenting views in climate science are also one of the main examples for the last type of overemphasis I want to mention: "inappropriate dissent". Traditionally, dissenting opinions have been defended and valued in philosophy, both for epistemic and ethical reasons. While perhaps best the known classical proponent of such a position might be John Stuart Mill, the value of dissent has also been highlighted in social epistemology, with authors like Helen Longino (1990), who, in her concept of transformative criticism (ibid., 76) emphasizes the importance of diversity and mutual criticism for the objectivity of science. Recently, however, it has been argued that some dissent can actually be quite harmful to science and society. While there is a philosophical debate about how inappropriate dissent is supposed to be identified and how fruitful it would be to target this dissent (Leuschner 2018; Melo-Martín and Intemann 2018), there are also clear cases where dissent is widely accepted as problematic: for example, the opposition to the fact that cigarette smoke is carcinogenic (Oreskes and Conway 2012), climate change dissent—where the extent or even the existence of anthropogenic global warming is denied (Biddle and Leuschner 2015)—, or creationists who deny the "history of the universe offered by the public system of knowledge" (Kitcher 2011, 155). What is especially interesting in the context of imbalance in the SoR, is that proponents of inappropriate dissent sometimes appeal to a lack of balance to further their agenda. Consider, for example, the argumentation of Oreskes and Conway (2012) concerning the strategies employed by the tobacco industry:

> The industry's position was that there was "no proof" that tobacco was bad, and they fostered that position by manufacturing a "debate", convincing the mass media that responsible journalists had an obligation to present "both sides of it". [...] The industry did not leave it to journalists to seek out "all the facts". They made sure they got them. The so-called balance campaign involved aggressive dissemination and promotion to editors and publishers of "information" that supported the industry's position. (ibid., 16)

It is important to ask what distinguishes these manipulative strategies from legitimate calls to diversify the SoR—which, at least partly, comes down to the question of what is the difference between problematic and unproblematic cases of imbalance that I address in section 3.2. However, dissent itself does not quite fit the concept of imbalance in the SoR introduced above: imbalance is described as there being too much or not enough research of particular kind; dissent, however, is about who holds and promotes what opinion in a scientific debate. Therefore, in the following, examples of inappropriate dissent itself will not play a major role in the analysis of imbalance in the SoR, safe for cases such as with the pause in anthropogenic global warming, where dissent leads to other forms of imbalance.

3.1.3 Categories of Imbalance

In the introduction to the last section, I have introduced different types of imbalance—lacunae, overabundance, and the various subtypes—as a way to guide through the list of examples. However, these are not the only systematical differences that can be ascribed to the phenomena. When we are concerned about there being too much or not enough of some kind of research, another question to ask is what we mean by "kind". In the cases mentioned above, there are at least three distinct categories: the major part are phenomena which are concerned with the distribution of **research topics**—think of profitable drugs versus alternative approaches in the case of unpatentable research, or already well-known parts of the genome versus disease-related but unexplored proteins in the case of ignored genes. A second way to categorize research would be by result: publication bias is about the idea that a scientific project is less likely to lead to a publication if it concludes in negative rather than positive findings. Similarly, also the debate described under the headline of evidential standards of EBM can be interpreted as being about what kind of evidence and thus what **research results** are produced[4]. That is, only results which constitute evidence derived from randomized controlled trials supposedly constitute relevant contributions. The case air-monitoring standards provides an example for the last kind of categorization I will highlight here[5]: there, the authors criticized that the evidential standards concerning air regulation might

4. This, however, is not the only way to describe this: it could also be described as being about the methodology used, or the source of the data under investigation.

5. The phenomenon of normatively inappropriate dissent would certainly constitute a fourth kind, where the relevant category would be an imbalance in the representation of positions within a debate. However, as I have explained above, dissent is a special case that I will not focus in this dissertation.

exclude the contributions of activists and concerned citizens. This, however, is neither a question of topic nor of result, but of who are the **research contributors**. This last category is especially interesting because it shows that imbalance is not just about the informational content of the SoR, that is, about the results of the research, but also about the scientific process itself and who can participate in it.[6]

While in these last cases, a lack of influence seems to be the problem, it is an essential part of most case-descriptions that problematic influences originating with particular groups of people are made responsible for the imbalance. In most examples, the connections between science and commercial influences are at least of some importance, while the concrete way in which these financial interests interact with research distribution are manifold. In the case of suppression, for example, financially invested parties themselves might directly prevent certain kinds of research from being made public, but in the cases of regulatory paradigms and air monitoring standards, the complex interaction of industry influences and policy regulations is presented as the source of imbalance. In the example of the pause in global warming, while of course there might be financial motivations in play, the criticisms of climate science that motivated the research focused on this topic could also have a more genuinely political background. Lastly, with the case of ignored genes, there is also at least one example where the motivations and limitations of scientists themselves—and only some of them based in financial interests—are cited as reasons for the distribution of research being considered problematic.

We can also distinguish the different types of decisions throughout the scientific process that can lead to imbalance. This is particularly clear in the case of publication bias: the decision of researchers to pursue a project and later submit a paper on it can influence what kind of research is published, but also the decision of reviewers and editors to accept a scientific article for publication. Other important avenues of influence are what projects receive funding—as is the problem in, e.g., the cases of unpatentable research or neglected diseases—or what should be the scientific standards for accepting some research as relevant for the SoR—as in the case of the debate about

6. I will revisit this point in subsection 6.2.2 when discussing participatory epistemic injustice as a reason for imbalance.

EBM or air monitoring.

The last aspect I want to discuss here is the area or field of research where examples of imbalance in the SoR have been described: strikingly, all the examples above are either connected to biomedical research, or to environmental science. This, however, is no mere sampling error or coincidence: the discussion of imbalance within the sciences itself has been almost exclusively focused on these fields, as have the meta-scientific disciplines—that is, mostly sociology and philosophy of science—dealing with the issues at hand. This should come as no surprise. Consider this description of the history of science ethics since the 1950s:

> Despite two articles in Science that argued the need for an ethical code for scientists, little was done by professional societies to address the need for guidelines for research conduct until the mid-1980s or later. For example, it was not until 1991 that the American Physical Society issued their first statement of ethical guidelines, (and those guidelines dealt exclusively with matters of research ethics.)
>
> In contrast, most engineering societies and at least one scientific society, the American Chemical Society (ACS), issued codes and guidelines for professional responsibility for a half century and more. Those statements had set forth norms of professional responsibility for public health and safety. Health care professionals such as nurses, physicians, and physical therapists had all established ethical norms for practice and to varying degrees educated new practitioners about their responsibilities. (Whitbeck 1995, 322)

While issues connected to the ethics of science have now become of interest on a broad basis, the discussion originates in the areas mentioned in the second part of the quote, and also the research connected to them came into the focus of ethical discussion sooner and more extensively than other areas of science. This is also apparent in the institutional backing of inquiries into these problems, with, e.g., the very substantial aforementioned reports on dissemination bias (Song et al. 2000; Song et al. 2010) coming out of the Health Technology Assessment (HTA) programme of the National Institute for Health Research (NIHR) in the UK, or reports on financial conflicts in medicine—where conflicts of interest in medical research only appear as one issue in the medical system at large—such as Lo and Field 2009, which was produced by the Institute of Medicine, part of the National Academies who consider themselves "Advisers to the Nation on Science, Engineering and Medicine" (ibid., 5).

This tight connection of science ethics with medicine but also with health and the environment in general is no contingent historical development either. Quite obviously, research that affects "public health and safety" directly, much more easily raises questions about the ethical evaluation of this research than, for example, basic research in theoretical physics. While the more genuinely epistemological questions connected to imbalances in the SoR should be similarly interesting in all areas of science, practical and ethical considerations—which make up the bulk of what I will discuss in second part of this dissertation—are much more apparent in these fields, where benefits and dangers to society are very direct. The cases from these areas thus allow me to clearly illustrate the problems with scientific research I really want to address. They are also representative of the discussions and philosophy and beyond which I want to connect to. Therefore, the perhaps somewhat narrow range of examples in terms of area of research should be no obstacle for the arguments presented here.

3.2 The Concept of Imbalance

In this section, I will give a tentative first answer to the question of what constitutes the concept of imbalance that connects the cases I have described in this chapter. In section 3.1, I have already mentioned a few characteristics of imbalance, which the following definition combines:

> **Def:** An **imbalance in the state of research** is a structural problem with the distribution of research which is constituted by there either being not enough (lacuna) or too much (overabundance) research of a particular kind, and which does not require that any individual scientific finding is problematic.

Note that in the context of relevance, my use of the term imbalance has been confined to particular SoRs; that is, to the research on one topic, or even more specifically, on the scientific information connected to one question of interest or decision to be made. It might be criticized that this focus is too narrow: what seems to be a problematic overemphasis of one line of research when we focus on a single decision problem may be justified if we consider its relevance for all possible questions of interest and applications. This is amplified if we consider not only relevance but also significance: surely, some questions are more pressing than others, so a neglect of relevant information in one area may be an acceptable compromise when the resources are put to a more urgent or important use elsewhere. While the concept of imbalance in this dissertation does focus on specific topics, it is not impossible to use it to analyze these overarching issues as well. Think back to the discussion of relevance in section 2.3: We can zoom in on any node in a network of inferences (cf. Figure 2.6) and discuss the research relevant to that specific question. For example, instead of asking "What is the SoR connected to the approval of this drug?" we can go further up the chain of inference and ask "What is the SoR connected to the safety of this drug?" or "What is the SoR connected to the carcinogenity of this drug?". But similarly, we can also move to a more general level, asking, for example, not only what is relevant for the decision about one particular drug x, but concerning all drug approval procedures. In the end, we could consider the all-encompassing issue of what research is relevant to any question we might have, and any decision-problem we might face. If we were to draw a graph of this SoR and the relations between evidence and relevant questions, it would, of course, be enormous both in terms of nodes and their interconnections—but it would not be qualitatively different from the graphs we have considered so far. Also, already on the lower levels, we have to make comparisons of relative relevance: is knowing about the safety of drug x more relevant to our pragmatic-epistemic goals than knowing about its effectiveness? Of course, weighing the relative importance, and in the end, the relative significance of research on more general levels becomes increasingly more complicated, and we might doubt that it even can be done systematically. In principle, however, we do not need to restrict ourselves to a specific topic when using the concept of imbalance in the SoR I have described so far. The reason why I do confine my analysis to particular decisions and questions lies in the examples under discussion: in all of them, the criticism is not about the overall importance of, for example, some line of inquiry all things considered, but instead about an imbalance that becomes apparent when weighed against some concrete alternatives. That is, the explanandum in this dissertation is not imbalance in the SoR as the entirety of facts about all of relevant scientific research, but about imbalance as a pro tanto problem.

In the following, I will turn towards another issue. I have described different ways in which imbalance can be problematic: research may not be published, not considered relevant, or not available at all; research might be wasteful, or influence the perceptions of the SoRs audience in an undesirable way. But why exactly are these cases of lacunae or overabundance to be considered problems? In subsection 3.2.1, I will examine two intuitive interpretations of balance—one based on a uniform distribution of research, one on freedom of science— but end up discarding both of them. Then, in subsection 3.2.2, I will discuss connections between imbalance and other, related concepts.

3.2.1 A Balanced SoR?

The word "imbalance" itself is a negative term; it suggests that there is such a thing as balance, from which the case under discussion deviates. But how is a balanced SoR supposed to look?

Balance as equal distribution. We could very simply demand that research should follow an equal distribution. However, it can easily be shown that this is neither possible, nor is it even a desirable ideal: one of the major ways in which the imbalance of the SoR was criticized in the examples given in the last section, was in terms of the topics under investigation. But already in the last chapter, I relied on Kitcher's argument about significant truths (cf. Kitcher 2011, 106) to claim that about any given object of research, we can ask an infinite amount of questions. Given that science is always limited by the amount of resources such as money, manpower, and time itself, it is simply impossible to pursue every possible topic and all research questions to *any* degree. And while refraining from engaging in any research whatsoever would result in an equal distribution, this hardly would be preferable over an imbalanced SoR. Also, even if we can pursue each of a particular set of research topics to the same degree, it is not clear why we should want to. As hinted at in the discussion of waste research, certain projects might seem more fruitful than the alternatives, and if this is the case, it would rather be problematic too spend equal resources on all of them.

To make things worse, not only can't we hope for an equal distribution of research given a particular category—say, positive versus negative findings or research done by the industry versus research produced by other parties. It is not even clear which categories should be considered for being in balance or not in the first place. We can characterize virtually any SoR as balanced or imbalanced by setting up the right categories, as ridiculous as that would seem: for example, we could ask if there is an overabundance of publications whose conclusion sections have less than one hundred words, or whose first author's surnames starts with a "D". While one may argue that it is quite obviously weird to consider categories like these, the categories that actually are considered in the literature are manifold—as I have shown in the last section. Also, some questions about science that will seem perfectly reasonable concerns for us now —for example, how inclusive it is in terms of gender distribution among researchers contributing to the SoR—might have seemed very weird in earlier times as well. Thus, we can never hope to find a SoR which is balanced in all categories, either; we need to make some selection of *relevant* categories in order to be able to say something informative about the SoR. It thus seems that imbalance, simply in terms of an unequal distribution of research, is ubiquitous, unavoidable, and in some sense even desirable. If so, what distinguishes problematic cases such as the ones presented earlier in this chapter?

Balance as independence. In the introduction, I have connected imbalance to worries about science as an independent source of information. Can we understand imbalance in terms of independence, that is as cases where some problematic influence on the SoR leads to more research of a particular kind or less research of another? This line of inquiry immediately leads to the question of "more" or "less" than what? In the meta-sciences, independence of science is a term rarely used, and even less often clearly defined. It is, however, mentioned in research ethics, appearing, among other things, in declarations such as "The European Code of Conduct for Research Integrity" issued by the European Science Foundation (ESF), where it appears one of the principles that make up scientific integrity:

> *Impartiality* and *independence* from commissioning or interested parties, from

> ideological or political pressure groups, and from economic or financial interests. (European Science Foundation 2011, 13)

Here, independence appears as the opposite of external influences which might lead to research misconduct. In philosophy, a similar concept which demands that there be no outside control of science has been discussed with the ideal of freedom of science (Wilholt 2010, 2012). While independence above appears as a duty for researchers, the concept of freedom puts more emphasis on the right of scientists to be in control over various decisions throughout the research process. We could thus be tempted to think of problematic cases of imbalance as issues with scientific freedom, where the "too much" or "not enough" of some type of research is relative to what would have been the SoR if science, in the determination of its goals, projects, and standards had been left to its own devices, protected from any outside interference. However, this ideal is far from universally well received. Daniel Sarewitz (2016) goes so far as to foretell the doom of modern science when it is followed too closely:

> Advancing according to its own logic, much of science has lost sight of the better world it is supposed to help create. Shielded from accountability to anything outside of itself, the "free play of free intellects" begins to seem like little more than a cover for indifference and irresponsibility. (ibid., 40)

While we might not share his gloomy outlook on science—or his analysis which attributes most of what is wrong with modern science to its internal logic, and in which he neglects to emphasize, for example, the effects of commercialization—also more nuanced analyses criticize the idea of the freedom of science as a carte blanche for science. Torsten Wilholt (2010), for example, highlights two justifications for the freedom of science: the potential epistemic benefits and the role of free science in a democratic society. But he also points out that one must differentiate what freedom of science is supposed to apply to—for example, freedom of the ends versus freedom of the means, or the freedom of individual scientists versus the freedom of science as a community (cf. ibid., 175) and that for each specific case, the arguments for scientific freedom "must be weighed against competing societal interests and values" (cf. ibid., 180).

It may be possible to focus on only those imbalances that actually do constitute a problem with scientific freedom, but not all the examples presented in subsection 3.1.3 fit this category. While I have admitted there that the major part of problematic influences that are perceived as the cause of imbalance are attributed to external factors, I have also already pointed out that some imbalance, where,—similar to Sarewitz' criticism—is described as resulting from internal problems, as in the case with the example of ignored genes.

As the examples show, intuitive understandings of the individual problems can substantially diverge: it may be about missing out on fruitful research or ignoring important results, about including contributions by certain groups of people, or about misleading the public and wasting the resources it provides for science. There simply is no single criterion for what makes these imbalances in the SoR problematic, and no simple answer to the question of what makes a balanced distribution.

3.2.2 Conflicts of Interest, Bias & Objectivity

Independence and freedom are not the only concepts which are tightly connected to the problem of imbalance in the SoR:

Imbalance as conflict of interest. Some of the examples discussed come from a discussion in research ethics that instead uses the term "conflict of interest": financial conflicts of interests might be suspected as a reason behind examples like unpatentable research, neglected diseases, suppression, or multiple trials with predictable outcomes. In one of the most comprehensive analyses of conflicts of interest in medicine, it is defined as follows:

> A conflict of interest is a set of circumstances that creates a risk that professional judgment or actions regarding a primary interest will be unduly influenced by a secondary interest. (Lo and Field 2009, 46)

Obviously, this definition itself leaves open what the "primary interest" of medicine that might be threatened by, e.g., a vested financial interest in a particular outcome of medical research, actually is. The authors, therefore, also provide some more specific ideas:

> Primary interests include promoting and protecting the integrity of research, the welfare of patients, and the quality of medical education. (ibid.)

While this makes things somewhat more clear, the only reference to research in particular is its "integrity". But research integrity itself is a term that, while widely used, often lacks a clear definition. In one of the rare cases that it is defined, however, one of the elements considered constitutive of integrity is, once again, independence in terms of avoiding external influences (European Science Foundation 2011, 7). Not only does this beg the same questions about independence outlined above, but it also is very close to the initial idea of conflicts of interest itself, and thus such attempts at defining the concept often appear somewhat circular.

Imbalance as bias. More useful to the purpose of this dissertation, perhaps, is the concept of "bias". With publication bias, there is at least one examples for imbalance in the SoR that constitutes a form of bias. Even etymologically, the word "bias" is related to imbalance; one of its earlier uses in the English language is connected to the game bowls, "where it was a technical term used in reference to balls made with a greater weight on one side (1560s), causing them to curve toward one side" (*Bias | Origin and Meaning of Bias,* n.d.) and thus refers to an object that is out of balance in a physical sense. Also, another thing shared between the examples of imbalance and the concept of bias is the idea that the problem does not only occur randomly or as the result of an individual, idiosyncratic decision of a particular actor, but here and there it involves some more or less stable tendency towards some kind of results which is the problem. More than just systematic, many imbalances of the SoR are connected to *systemic* biases: opposed to the idea of personal biases, a systemic influence is not simply a regular tendency—which might also apply to a single individual—, but refers to cases where the—perhaps institutional—context of a type of decision constrains it in a way that makes specific outcomes much more likely.

Why then, it may be asked, do I not generally use "bias" instead of "imbalance", given that the former is already an established term in the sciences? Firstly, bias-concepts, as in the example of publication bias, refer to some specific mechanism, while imbalance is an umbrella term for the different mechanisms described in the examples. Also, while—at least in connection with scientific research—bias normally refers to an influence on a decision, imbalance describes the very *state* of the research. This is important because not all the examples highlight a problem with an influence on the SoR; the cases that constitute examples of what I have called overemphasis, for example, are explained in terms of the biasing effect on public perception that the SoR can have—and not the other way around. We can thus understand imbalance in the SoR as a type of problem where the lack or the overabundance of some kind of research is problematic either because it is the result of some bias, or because leads it leads to biases further down the line.

Balance as scientific objectivity. Still, neither independence, nor freedom, conflicts of interest, or bias provide a clear evaluative criterion for assessing whether an imbalance is problematic. What can we say about the requirements for such a criterion? For one, it both needs to be able to account for epistemic as well as more practical and ethical considerations. In some cases, for example when referring to instruments used to measure some outcome—perhaps in a scientific experiment—, we might say that bias is a systematic deviation from some true value. But truth is not necessarily the standard a bias deviates from. In cases of imbalance in the SoR, it is generally difficult to say how they might constitute a problem with truth: imbalance is not about the individual result but the distribution of research, after all—and how can a distribution deviate from the truth?[7] But also aside from imbalance, there are entirely different kinds of bias: Would a biased sentence delivered by a biased court of law have to be *untrue*, or, being a matter of justice, would it not rather be unjust? Examples like the air-monitoring standards were described in terms of who can contribute to science. In terms of diversity, this might also have epistemic effects, but it mainly seems to be about fairness towards the various social groups invested in the results of the research; i.e., it is an ethical concern. There is one concept connected to bias—or rather: radically opposed to it—which can accommodate this variety of perspectives and is also central to philosophy: scientific objectivity. Reiss and Sprenger (2017) give the following explanation of the concept:

> Scientific objectivity is a characteristic of scientific claims, methods and results. It expresses the idea that the claims, methods and results of science are not,

7. I will follow up on this question in section 5.1.

> or should not be influenced by particular perspectives, value commitments, community bias or personal interests, to name a few relevant factors. Objectivity is often considered as an ideal for scientific inquiry, as a good reason for valuing scientific knowledge, and as the basis of the authority of science in society. (ibid.)

While the breadth of this description makes it at least intuitively plausible that we could try to assess all the examples given in this chapter in terms of objectivity, it also indicates that, perhaps, objectivity might not be one monolithic concept, but can be understood in various ways. Douglas (2004), for example, distinguishes eight forms of objectivity in three different, reportedly irreducible modes: from manipulable objectivity$_1$—which is about "how reliably and with what precision we can intervene in the world"(ibid., 457) to interactive objectivity$_3$, which takes up Longino's account of transformative criticism (cf. Longino 1990, 76). This variety is even more impressive considering that Douglas restricts herself to objectivity in terms of scientific processes—as opposed to, for example, the products of science—and also does not take into account the historical meanings of objectivity as described in detail by Daston and Galison (2010).

How should we begin to apply such an expansive and heterogeneous concept to the cases of imbalance in the SoR? Instead of going through each of the meanings of objectivity ever described and checking it against all the different types of examples presented in this chapter, in the second part of the dissertation I will present a normative framework for assessing imbalance which takes its starting point in the concept of public epistemic trustworthiness[8]. But why should we look towards trust when concerned with objectivity? Describing the perspective of instrumentalism in the debate about objectivity, Reiss and Sprenger (2017) provide the following rationale:

> We want scientific objectivity because and to the extent that we want to be able to trust scientists, their results and recommendations. One possible lesson

8. You may wonder why I chose to focus on public trustworthiness, given that I am generally interested in the SoR also as a resource for individual decisions—not only for public policy. But the concept developed in chapter 4 can easily be transferred from the public as a whole to individual members; the level of the public, on the contrary, introduces additional requirements in terms social responsibility, which would not be captured by an account in terms of individual trust-relationships; thus, I focus on the more demanding concept.

> to draw from the fairly poor success record of the proposed conceptions of scientific objectivity is that these conceptions have the logical order of the ideas mistaken. They look at some privileged feature of science, define this feature as "objectivity-making" and then leave the issue of whether or not the feature also promotes trust to fate. The obvious alternative is to reverse that order, start with what we want and then look for features that might promote the thing in which we are ultimately interested. (ibid.)

I will not commit to an outright instrumentalist position myself, claiming "that anything that stands in the right kind of causal relation with public trust will count as an objective feature of science" (ibid.): not all elements constitutive of epistemic trustworthiness contribute to the objectivity of science.[9] However, starting my analysis with the concept of trustworthiness provides a perspective from which, by analyzing the different requirements for trust, I will systematically develop the different criteria than can then be applied to the examples of imbalance in the SoR. Also, I had declared it a central research goal that the analysis of imbalance provided in this dissertation should contribute to our understanding of how the information provided by science can and should be used in policy-making and the deliberations of individuals. Building a conceptual analysis of imbalance on a perspective of public epistemic trustworthiness ensures that the following discussion is focused on this connection between scientific objectivity and society at large.

There is one more complication concerning the relation of trustworthiness and imbalance that needs to be mentioned: the existing discussion of objectivity does not quite fit imbalance as introduced here, because imbalance refers to problems with the distribution of various types of scientific results. Both in Reiss and Sprenger (2017) as well as in Douglas (2004), however, objectivity is considered only in terms of individual phenomena: "For example, what does it mean to say that a particular experiment produced an objective result, solely in terms of the interaction between the human experimenters, their equipment, and their results?" (ibid., 456) Even concerning

9. Consider the requirement of benevolence or good-will I describe in section 4.1: while it seems likely that some *y* being positively disposed towards *x* increases *y*'s trustworthiness, without further qualification it would indicate that they are in fact biased towards *x*, which intuitively does at least not seem to increase objectivity, if not diminish it.

objectivity$_3$—which is about “the social processes in knowledge production” (ibid., 461) objectivity is still supposed to be about particular processes of justification and what it means “to claim that the end result is objective” (ibid.). While I have already mentioned that mechanisms like “value-laden blind spots” (Bueter 2015, 18) in theory assessment can be used to show connections between imbalances in the SoR and the justification of individual results, the main claim I investigate in this dissertation is that the distribution of research can be problematic even if individual propositions are not. This apparent disconnect between objectivity and imbalance I perceive not just as a hurdle, but as an opportunity: applying a framework based on objectivity as a precondition of trustworthiness to cases of imbalance in the SoR will not only provide a more systematic analysis of these cases; by discussing how the various concepts connected to this form of objectivity must be interpreted to accommodate imbalance, I will also advance the theoretical discussion about objectivity.

I.: Lacunae

Phenomenon	Type	Neglected	Favored	Category
Regulatory paradigms (Frickel et al. 2010)	Undone science	Assessing the danger of chlorides as a class	Ad-hoc assessment of the toxicity of individual compounds	Research topic
Unpatentable research (Brown 2008)	Undone science	Alternative approaches	Profitable drugs	Research topic
Neglected diseases (Reiss and Kitcher 2009)	Undone science	Diseases of the less-affluent	Diseases of the wealthy	Research topic
Ignored genes (Edwards et al. 2011)	Undone science	Large parts of the human genome	Few well-researched genes	Research topic
Evidential standards of EBM (Stegenga 2011)	Excluded research	Mechanistic/casuistic evidence	RCTs & meta-analysis	Research results
Air-monitoring standards (Frickel et al. 2010)	Excluded research	Community-based research	Expert research	Research contributors
Publication bias (Song et al. 2010)	Inaccessible research	Negative results	Positive results	Research results
Suppression (Chan et al. 2014)	Inaccessible research	Findings opposed to interests	Findings favoring interests	Research results

II.: Overabundance

Phenomenon	Type	Neglected	Favored	Category
Multiple trials with predictable outcomes (Sismondo 2008)	Research waste & Overemphasis	—	Research on blockbuster drugs	Research topic
“Pause” in global warming (Lewandowsky, Risbey, and Oreskes 2016)	Overemphasis	Other climate research	Research concerning hiatus	Research topic

Table 3.1: List of examples for imbalance in the SoR

Part II

The Normative Background

Chapter 4

Imbalance and Trustworthiness

In the previous chapter, I have proposed a framework of criteria that can be used to explain why we intuitively think of the examples given in section 3.1 as problematic. The construction of this framework was supposed to systematically proceed from an analysis of scientific trustworthiness. Trustworthiness is not a new issue in discussions about science and its relation to the public. In the editorial introduction to a special issue about trustworthy research, Caroline Whitbeck (1995) describes philosophical analyses of trust in science as the latest development in an ongoing discourse about research integrity that originated in the mid-twentieth century. After an initial period, in which the acknowledgment of problems with the ethical conduct of science was an isolated phenomenon, a second phase, starting in the nineteen-eighties, saw the formulation of ethical codes in various scientific disciplines. Confronted with cases of flagrant scientific malpractice, these guidelines introduced legal and semi-legal frameworks, mostly geared towards preventing the trinity of research misconduct: fabrication, falsification, and plagiarism. Whitbeck locates the discussion about trust and trustworthiness in science within a third phase, which is constituted by a move from legalistic approaches focused on the fraudulent behavior of individuals towards research ethics under a broader ethical perspective. Since then, philosophers such as John Hardwig (1991), David Resnik (2011), and Torsten Wilholt (2013) have discussed the trustworthiness of science, drawing on general philosophical accounts of trust provided by central figures such as Anette Baier (1986) and Russell Hardin

(1996).

One major development in philosophy of science in these last decades has been a growing emphasis on the social dimension of science. This is connected to the philosophical framework of social epistemology, which emphasizes the role of testimony—and thus of credibility and trust—in epistemic processes. In this context, epistemic trust—i.e., trusting someone "in her capacity as a provider of information" (Wilholt 2013, 233)—is often seen as a precondition for the success of research, especially in collaborative settings (Rolin 2015; Wilholt 2016), which have become more and more prevalent in science. However, discussions of epistemic trust and trustworthiness have been focused almost exclusively on the functions and preconditions of trust in the justification and dissemination of individual scientific findings. Especially when we recognize the role of trustworthiness not only as a requirement for collaboration between scientists but also as a precondition for its usefulness as a source of information for society (Whyte and Crease 2010), this is not the whole story.

How is scientific trustworthiness related to science for policy and individual decision-making? In the previous parts of this dissertation, I have suggested that only a comprehensive picture of available findings concerning different decision options and their relative confirmation and support constitutes the adequate level of scientific information: the SoR. Consequently, also when we assess the trustworthiness of science, the SoR is what we need to be concerned about most.

But what, generally, is trustworthiness? In section 4.1, I will provide a definition of the concept in terms of a three-place relation and three requirements demanded of a trustworthy party. In section 4.2 I will apply this general definition to imbalance in the SoR as a source of information for the public; first, I will discuss the special requirements of trustworthiness on science as a provider of information. Next, I will state more precisely what should be the relata in this trust relationship: what is the object of trust, who trusts, and who is trusted? In section 4.3, I review the framework presented in this chapter and relate it to the other parts of the normative background of imbalance in the SoR.

4.1 A Definition of Trustworthiness

The explanandum in this section is trustworthiness, which is a concept both more and less ambitious than some related notions: *actual trust* refers to any occurrence of someone trusting someone else. It does not require that the trusted is trustworthy, and it may be influenced by all kinds of, e.g., psychological factors. Also, actual trust can come in both cognitive and non-cognitive forms (Baier 1986; Becker 1996). While this is an interesting topic for empirical investigations of trust, in this part of the dissertation, I aim to provide normative criteria for evaluating imbalance in the SoR, not with descriptive concepts. *Warranted trust,* on the other hand, is indeed normative: Gürol Irzik and Faik Kurtulmus define warranted trust as "trust with good grounds invested in those scientists with the required qualities for being trustworthy" (Irzik and Kurtulmus 2018, 2). This is, however, too demanding for my purposes: I am concerned with evaluating problematic distributions of research; warranted trust, however, does not only demand trustworthiness on the side of science, but also the awareness of this trustworthiness on the side of the trusting party.[1] What, then, is trustworthiness?

Trustworthiness, reliance, and competence. I consider trust to be a special form of reliance. Reliance is a three-part relation: Party *A* relies on party *B* for some *x*. For *B* to count as reliable, they will have to be competent in performing *x*. But this is not enough: not only must B be able to perform x, but they also must actually do so when called upon. If I am going to trust a bicycle repair shop with replacing the brakes on my bike, not only will I have to assume that the mechanic knows what to do, but also, for example, that they will not cheat me by using inferior parts or claiming that more parts need replacing than actually do.

In some cases of reliance, *A* will actively seek control over *B*'s performance of *x*. Relying on the repair shop, I could try to safeguard my interests through legal measures, such as the ones guaranteed to me in a written contract. Alternatively, I might want to monitor *B*'s performance of *x*, in so far as I am able: I could ask to be

1. While the epistemic situation of the public as the trusting party does not primarily concern me in this dissertation, I will touch upon this issue when concerned with imbalance as a communication problem—cf. subsection 5.1.1—and epistemic trust injustices—cf. subsection 6.2.2.

in the repair shop while they do the repairs, to see if they are done in a professional manner, to check that the bill they send me reflects the actual labor hours, and so on and so forth.

But as the proverbial "trust is good, control is better" implies: trust and control are disparate things; contrary to other forms of reliance, for *A* to actively ensure *B*'s performance of *x* is out of place in a trust-relationship. It would seem strange to claim that I trust my spouse not to have an affair on their working trip if I then send a private investigator after them to ensure the very same.

When the truster (*A*) relies on the trusted (*B*) for *x*, *A* forfeits control to the extent of how much one trusts the other; *A* incurs risks by granting—using Annette Baier's term—some discretionary power over *x* to *B*. Note that relinquishing control is not only a precondition for but also the major advantage of trust-relationships over other forms of reliance: Making sure that the other party can actually be relied upon can be time-consuming, troublesome, or outright impossible, especially if I rely on someone to perform a very specialized task.

Trustworthiness and benevolence. Trustworthiness, then, in contrast with reliability, does not only require competence on *B*'s part, but another quality which makes it possible for *A* to rely on *B* for *x*—without actively ensuring *x* through control of *B*'s activities[2]: "When I trust another, I depend on her good will toward me." (Baier 1986, 235) If I am at someone's mercy, I will need to assume that they mean me no harm and have my best interests in mind if I am supposed to trust them. But the addition of "good will", as a benevolent attitude from one individual towards another, is still not quite enough.

Trustworthiness and integrity. Even if the trusted is capable and well-disposed towards the trusting party, *A* might still question if *B* has the appropriate attitude towards performing x; after all, *B* might be lazy, negligent, or generally adhere to views on the subject which are incompatible with what A considers to be the proper moral or professional standards. Thus, alternatively or complementary to good will or benevolence, we can posit integrity as a requirement for trustworthiness. Integrity can be understood in terms of principles or dispositions—that guide *B*'s discretion

2. I follow the exposition in Mayer, Davis, and Schoorman (1995).

in regards to performing *x*—being shared or well-regarded by *A*[3]. If I am to trust a candidate to represent me in parliament, this might presuppose shared political views; if I am to regard the referee in my football game as trustworthy in giving their rulings, I might require them to adhere to a principle of impartiality—even if I myself am not impartial. We can now define trustworthiness:

Def.: *B* is worthy of *A*'s trust concerning *x* iff

1. *B* is able to perform *x* (**competence**)
2. The discretionary powers over *x* *A* grants to *B* are guided by
   - **(a)** *B*'s good will towards *A* (**benevolence**)
     and/or
   - **(b)** principles or dispositions concerning *x*, which are shared or well regarded by *A* (**integrity**)

This definition mirrors the three factors of trustworthiness—ability, benevolence, and integrity—identified by Mayer, Davis, and Schoorman (1995) in an analysis of multiple preexisting trust-concepts. Understanding the requirements above not as binary conditions but as *factors* influencing trustworthiness implies another important point: trustworthiness itself is a gradual concept, and, ideally, the scope of *x* as well as the amount of discretion *A* grants to *B* should reflect the degree of *B*'s trustworthiness.

In the next section, I will apply this definition to public epistemic trust in science, and highlight its particular features. Much of its specific character, I will argue, rests on the ideal of scientific freedom as both foundation of and limitation to the discretionary power invested in science.

3. This interpretation of integrity resembles Wilholt's (2013) analysis of epistemic trust in science as relying on "a shared sense of what the right attitude towards the aims of a collective epistemic enterprise is, and [...] the confidence that other participants in the enterprise actually display that attitude" (Wilholt 2013, 251). The author, however, does not define this requirement as integrity and also doesn't subscribe to a threefold model of trustworthiness in terms competence, benevolence, and integrity.

4.2 The Trustworthiness of Science

4.2.1 Requirements on the Public Epistemic Trustworthiness of Science

How do the requirements of competence, benevolence, and integrity apply to science being trusted as a producer of information? Competence as a requirement for trustworthiness is at least as important concerning the public's trust in science as it is in any other trust-relationship. In fact, the often enormous difference in expertise between the scientists and the recipients of scientific findings is crucial for understanding why it is so important that we can trust science: nobody, perhaps not even other scientists from similar fields, can hope to fully appreciate the work of some specialist in their respective niche. This is also true for other kinds of experts, but arguably most prevalent in science. If we thus cannot keep a check on the work of researchers, and we still depend on it, we can only put our trust in them. As a minimal explanation, for science to count as competent, it needs to be able to produce information that is reliable enough to improve the epistemic situation of members of the public given a specific context. However, I will not try to provide a detailed analysis of competence in this chapter. None of the examples of imbalance in the SoR seem to be about science being unable to provide helpful information in principle; on the contrary, they often are about cases where something could be done, but this option is ignored or neglected.

As for benevolence, while for science to be trustworthy it would be problematic if most of the people involved with science had a malevolent attitude towards the wider public, it is not immediately clear how this interpersonal component can be applied to science as a whole. I will, however, have more to say on this problem and good-will in subsection 4.2.4.

For now, I will focus on the special character of scientific integrity: In professional ethics, many ethical principles governing scientific research have been discussed; Resnik (2007), for example, lists a total of 14 of them. Among these, there are many principles of professional integrity which are based on general ethical considerations: for example, respect for colleagues, students, and research subjects are just instances of a more general respect for our fellow human beings—and, perhaps, animals. While these general norms are important for trust in science, there are also some principles which are of special relevance to science in its function as a producer of information for the public. Discussing these requires an investigation into the nature of the discretionary power granted to science by society: returning to Baier's account, if we want to explicate benevolence—and integrity, although Bayer herself does not use that term—, "we can look at various reasons we might have for wanting or accepting such closeness of those with power to harm us, and for confidence that they will not use this power." (Baier 1986, 235) What are the grounds for trusting science with the production of knowledge, instead of exerting a more stringent public control over the research process?

To some extent, trust is indispensable; modern science requires both a degree of specialization and trust-based collaboration between experts that makes direct oversight impossible. However, the discretion actually granted to science goes way beyond what is necessary. Consider the widely held ideal of scientific freedom: science is both heavily supported by public resources and enjoys a degree of autonomy that is unusually high compared to other publicly sponsored enterprises. This freedom constitutes a special kind of discretionary power which is granted to science by society. Therefore, justifications of scientific freedom, as reasons for granting said discretionary power, can be considered principles of integrity which should guide science in the use of that discretion. Following Resnik's principle of social responsibility, we can take this to imply "a contract that scientists have with society: scientists agree to help society in return for public support" (Resnik 2007, 48).[4] But how does scientific freedom help society?

There are at least two widespread and long-standing arguments for granting scientific freedom. The epistemological argument (Wilholt 2010, 175-177) rests on the premise that granting scientific freedom increases the epistemic productiveness

4. This concept of scientific integrity may diverge from the actual expectations held by the trusting party, i.e., the public. Instead of providing a descriptive, empirical account of what the public *does* expect of science, I here provide an account of what the public reasonably *should* expect from science, given the investment. A descriptive account—the problem of empirical investigation aside—would be difficult given that "the public" is not restricted to any particular society.

of science, be it due to an increase in the diversity of approaches, the heightened motivation of researchers or the more effective use of localized expert knowledge. This seems intuitively plausible, but not only does the strength of the argument depend on unsettled empirical questions; it is also debatable for another reason: even if we accept that scientific freedom leads to an increase in the output of science, it remains unclear if the findings thus produced are relevant to society. In the context of public trust and support, research needs to be valuable to a wider audience—be it in terms of answers to societal problems, technological advances, or simply epistemic interest—than just the residents of the metaphorical ivory towers. Much discussed ideas like well-ordered science (Kitcher 2011, 105), responsible research and innovation, public engagement, or citizen science emphasize this point.

A reason for the public to refrain from trying to enforce societal relevance directly is given by the political argument (Wilholt 2010, 177-179). Even for a democratic institution aiming at increasing the societal value of research, or so the idea, the attempt to dictate the scientific agenda would be self-defeating. The autonomy of science, it is presupposed here, plays an important role in the democratic process. Controlling science would impair its contribution to, for example, the formation of the political will and, therefore, undermine the democratic legitimation of the very political institutions exerting this control. This is because for science to provide reliable information and common ground to the concerned parties in a deliberative process, science needs to be independent of dominating influences originating with only some of them. Also, if the scientific process is partial to particular interests, this can promote the unjust distribution of epistemic resources.

These two arguments, however, do not translate to principles of integrity directly. In the following, I will explain the underlying criteria in more detail, disambiguate the two concepts expressed in the individual arguments, and lastly, in the case of the political argument, generalize from the specific context of scientific freedom to a broader principle in terms of epistemic justice.

The principle of epistemic productiveness. The first of the two criteria, which I will call epistemic productiveness, goes back to Wilholt's epistemological argument. There is some variance in his exact formulation of the reason that motivates this justification of scientific freedom:

- "What I will call the epistemological argument is a line of reasoning defined by its core premise, which can be formulated as follows: a principle of freedom of research creates **optimal conditions for our collective search for knowledge** [my emphasis] ." (ibid., 175)
- "In sum, the epistemological argument can plausibly show that under certain conditions—notably free communication and a functioning system of incentives—individualized freedom of research will lead to a diversity of scientific approaches that can be expected to **surpass the epistemic yield** [my emphasis] of centralized forms of research organization." (ibid., 177)
- "Note that the political argument, thus conceived, is not just a variety of the epistemological one. Its aim is not to **guarantee that the knowledge needed by the citizens is generated in as efficient a way as possible** [my emphasis]." (ibid., 178)

There are three important aspects that I want to highlight in these characterizations: firstly, productiveness is process oriented; it is about the way in which knowledge-production is organized. Secondly, it is not just about making the production of some knowledge possible, it is about increasing "yield", that is, about the efficiency of knowledge production, which also implies that productiveness is not a binary but a gradual concept: we can rank different forms of organizing knowledge production. And thirdly, productiveness is not just about any knowledge, but the "knowledge needed by the citizens", i.e., the questions we as a society want to see answered in "our collective search for knowledge". Especially the qualification "needed by the citizens" begs further explanation. First of all, it should be made clear that an idea of knowledge being useful is not supposed to favor applied over basic forms of research, or—perhaps even more dangerously—research that can be patented and monetized over less lucrative projects. In fact, the word "productiveness" was chosen to avoid the economic connotation attached to the more common "productivity". What *is* meant by this appeal to needs or utilities is very close to the notion of questions of interest used by Alvin Goldman (2003) in his framework for assessing veritistic value:

> There are three relevant senses of "interest." One measure of a question's interest is whether the agent actively finds it *interesting*, that is, has an aroused curiosity or concern about the question's answer. Such concern can arise from intrinsic fascination or from recognition of the potential practical value of knowing a correct answer. A second measure of interest is dispositional rather than occurrent. Many questions *would* be interesting to a person if he/she only thought of them, or considered them, although no such consideration has in fact occurred. Such dispositional interests should also be counted in assessing a question's "importance" or "significance" for that person. A third sense is more broadly dispositional: what would interest the agent if he or she knew certain facts. Students might take no active interest in a certain topic, yet such knowledge may be objectively *in* their interest because it is relevant to matters in which they do take an interest. (Ibid., 95)

For example, one might care to know about the changes in temperature during the last ice age—if only because of simple curiosity about the history of our planet—(actual interest); perhaps, while a patient or physician has never thought about an alternative to drugs in combating some disease, they would care if confronted with the idea (hypothetical interest). We can interpret the different kinds of interest analyzed by Goldman in terms of the notion of relevance I presented in section 2.3 (see Figure 4.1):

If one has either an actual or hypothetical interest in an answer to a higher-order question—for example, if a certain substance used by the manufacturing industry should be approved for use or banned from the market—, this also implies that one should have a subsidiary or instrumental interest in all the questions which—according to the normative principles one accepts—are relevant to this original question. In our example, this might imply an interest in the safety of the substance (but also, for example, in other questions such as if there are any alternatives that could replace the substance) which in turn might imply an interest in the question if this substance causes cancer (and once again, possibly a range of other safety concerns). At all these levels, the distribution of research in some SoR could be of more or less use to, or even

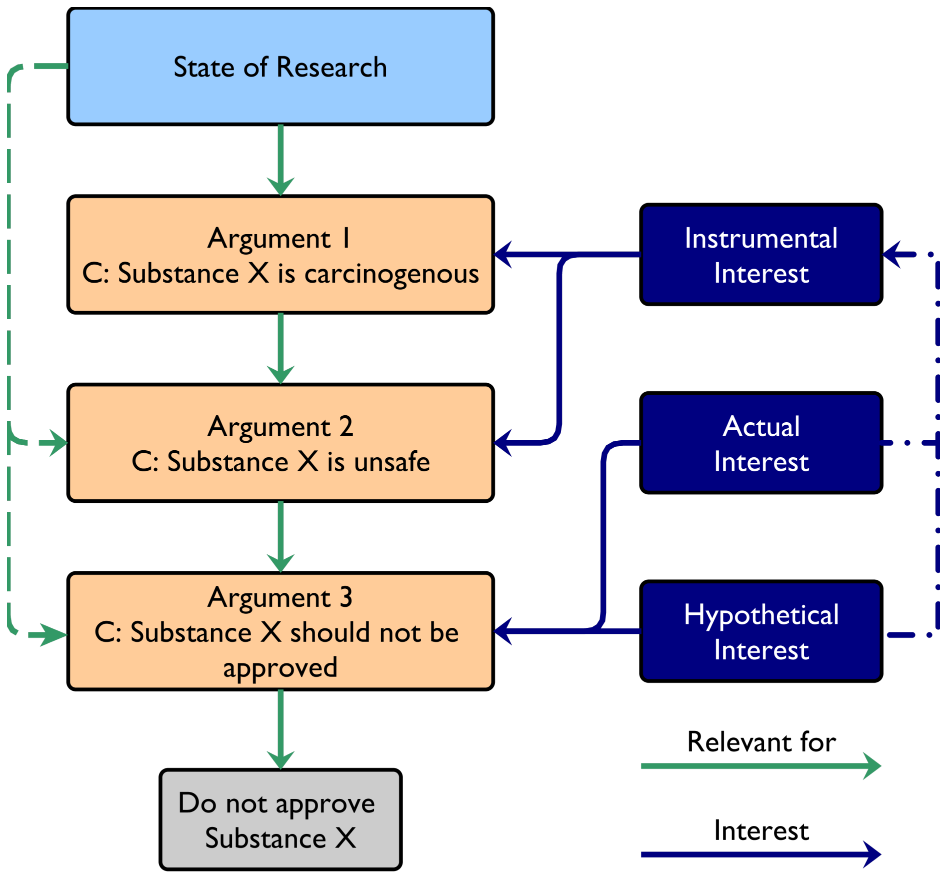

Figure 4.1: Relevance and Interest

hinder the attainment of relevant knowledge. Thus, an imbalance in the SoR might be seen as detrimental to the productiveness of science concerning actual, hypothetical, or instrumental interests.

The principle of epistemic justice. Highlighting the given interests of individuals also points towards the second criterion, which can serve as a basis for a principle of scientific integrity. As David Coady (2010) rightly laments, Goldman's account based on maximizing veritistic value in terms of epistemic interests "can be criticised for focussing exclusively on the amount (or average amount) of the fundamental good (in this case interesting true belief), rather than its just distribution" (ibid., 103). Also in the examples of imbalance discussed in chapter 3, some of the cases were described in terms of conflicting vested interests: consider the example of neglected diseases—where research focused on diseases of the wealthy as opposed to diseases of the less affluent—or the idea that the standards of the EBM movement might favor contributions by the industry over community based research. When the interests of one or more affected parties are favored in this way, imbalance in the SoR is about the second criterion of scientific integrity, which is based loosely on the political argument in Wilholt (2010):

> Its aim is not to guarantee that the knowledge needed by the citizens is generated in as efficient a way as possible. Its punch line is rather that it must be generated in ways that are independent from the major political powers, because no political power can have legitimate control over the generation and distribution of the same knowledge that serves as an input to the democratic process without thereby undermining whatever democratic legitimacy might otherwise justify its control over science. (ibid., 178)

Wilholt's argument, however, is very specifically tied to a self-contradiction that arises when we try to justify a restriction of scientific freedom with reference to democratic legitimization. He argues that even the most sensible ideas of unilateral control over the scientific process—control which would be legitimated by democratic procedures—can be self-defeating. Even if some party has been elected by democratic vote, for it to take direct control of the scientific agenda might interfere with a fundamental idea of democracy: for a democratic vote to be legitimate, the voters first need to be able to form a political will, depending on all the available information about the political issues at hand. If a political party is allowed to influence the available information in a particular direction, this interferes with the formation of the political will, and thus with the very basis of the procedures that could legitimate this control.

To arrive at a second criterion of scientific integrity, I expand upon this specific motivation for scientific freedom in two steps: Firstly, if private interests without legitimization dominate the SoR, this is not less but rather more concerning than democratic oversight of science. Therefore, the resulting principle should exclude any unilateral dominance over the processes that determine the SoR, not only those influences with democratic legitimization. Secondly, while the independence of scientific processes is a reason to grant science and its institutions some freedom, this is just the means to an end: ensuring that the contributions of science to society are fair to the legitimate interests of all its members. And while some of the negative effects of outside influences may be avoided by granting science trust and freedom, the interests of scientists and scientific institutions may also sometimes unduly favor their interests over those of other people, and thus lead to injustices: with cases of imbalance such as ignored genes and publication bias, I have mentioned two examples for cases where not only outside influences on science, but also the self-interest of scientists themselves can skew the SoR in problematic ways. In such cases, not scientific autonomy, but, on the contrary, independence from the private goals of scientists would have to be demanded. As I have argued in chapter 3, independence—and also alternatives like impartiality—are rather vague concepts. These values also bring with them the additional difficulty that they may be confused with a value-free ideal, which, at least when we consider issues such as agenda-setting, is not considered desirable by philosophers of science [5]. As also the criterion of productiveness suggests, the distribution of research should be motivated by the practical needs of society. How, then, should we formulate the second criterion? In chapter 6, I will broadly re-interpret the reason behind Wilholt's argument as a problem with a criterion of epistemic justice: no social group should be allowed to dominate the decisions shaping the SoR or unilaterally benefit from scientific research, as that would be unfair to the rest of society.

As an interim conclusion, I have specified the following as requirements for the public epistemic trustworthiness of science concerning the SoR: firstly, science needs to be competent in terms of being able to improve the epistemic situation of members of society concerning relevant questions. Secondly, science must not be hostile or malevolent towards the public. Thirdly, concerning the discretion granted to it by the public in producing the SoR, science needs to adhere to principles of integrity. Beyond general rules of good ethical conduct, science might reasonably be expected

5. In the recent discussion, the value-free ideal has been understood as the demand that scientists should refrain from using non-epistemic values as justifications for scientific claims. Even in this context, only few philosophers defend the ideal. While there is still some debate about the role of non-epistemic values in what may—relying on the perhaps somewhat antiquated tripartite division of the scientific process—be called the context of justification, there is, to my knowledge, no serious claim that in the contexts of discovery or application only epistemic considerations should be admissible.

to follow two special principles of scientific integrity, which I have derived from arguments for scientific freedom: as per the criterion of epistemic productiveness, science should aim to efficiently produce knowledge which is needed by members of the public. As per the criterion of epistemic justice, the decisions shaping the SoR should not be dominated by the interest of particular social groups, nor should they unilaterally benefit from scientific research.

In the next sections, I will explore the relata of the trust relationship in the context of the SoR: who trusts, who is trusted, and what are they trusted with?

4.2.2 *x*: The Object of Trust & Trust in Objects

The definition of the trust-relationship I have used in this chapter is based on Baier's concept of trust, which is about interpersonal trust: some individual *A* trusts some individual *B* concerning some object or activity *x*. In subsection 4.2.1, I have so far moved between talking about trusting scientists, trusting the SoR and trusting the process which shapes it. But are there not important differences between those trust-relationships? While I am concerned with imbalance in the SoR in this thesis, I do not analyze the trustworthiness *of* the SoR itself; i.e., this chapter is about the SoR as the object of trust *x* that the trusted party *B*—science—is trusted *with*. According to this analysis, trust in the SoR comes down to trust in those who produce it.

But is this the best way to conceive about trustworthiness in connection with imbalance in the SoR? Can we not also directly trust the SoR? Generally, does it make sense to talk about the trustworthiness of a product, an object itself, without this being short-hand for trusting the ones that produced it? In the following, I will argue that there is a cogent way of talking about trust in objects which is not just basic reliance, but which is also still not the same as the concept of trust that will be used in this dissertation.

In ordinary language, we can surely say that some object appears trustworthy—or not. Imagine a hiking trail in the mountains, with a narrow bridge that spans a deep ravine blocking your path. Does it not seem natural to wonder if we can trust, for example, that the wood will hold our weight and the way ahead is safe? But what does this estimation of safety actually come down to in the framework used in this chapter? I have located the main difference between trust and reliance in the relation between trust, discretion, and control. If we trust someone, we cannot exert full control over the service we trust the trusted party with; we have to grant them discretion, we have to put ourselves at their mercy and incur some risk. This latter part—exposing oneself to risk, but still implicitly or explicitly assuming that we will not be disappointed by that in which we trust—is also present in the example of the mountain bridge: we can say that we trust it, rather than rely on it, because we have to wager that we can safely pass over it although we are not quite sure that it will hold. If we were absolutely sure that the bridge is safe, it would not make sense to talk about trust in the first place. But we might still say that we will go over the ravine by ways of the bridge; i.e., that we rely on it. In this account of trusting the bridge, what separates trust from mere reliance is the acceptance of the uncertainty of success; the acceptance of risk. Can we reduce trust relationships to this alone? And would the same idea of trust that applies in the bridge example also work in the context of SoRs? Is it an ascription of trustworthiness, or rather an evaluation in terms of reliability?

Let us return to Baier's original text. The element of incurring risk is an important element of her account, but it is a special kind of risk: the risk of allowing oneself be vulnerable, in the sense that by trusting someone, we may be hurt if they bear ill will towards us. The following passage from Baier (1986) is perhaps the one in which the difference between relying and trusting stands out most:

> We all depend on one anothers' psychology in countless ways, but this is not yet to trust them. The trusting can be betrayed, or at least let down, and not just disappointed. (ibid., 235)

If the mountain bridge gives way under our feet, at that moment we may think that the wood has betrayed our trust, but surely we anthropomorphize the object in that moment. We may just as well say that we were let down or disappointed, and lose none of the content of the assertion. It was not the bridge's malevolence that caused our fall; not a failure to adhere to principles of integrity on its side. The bridge cannot have discretionary power, simply because it is no agent, but an object.

But are these interpersonal and emotional elements of trust so important in the context of SoRs? Could we not just analyze imbalance in the same terms as the mountain bridge, in terms of being uncertain about the reliability of the information provided by the SoR, of its usefulness to our endeavors? Does the possibility of the public being betrayed by science matter for this dissertation? Certainly, the question of whether a SoR is a reliable source of information is an important part of epistemic trustworthiness. However, by emphasizing the principles of integrity at work within the scientific process, by pointing towards the values shaping the SoR, we reveal more about the scientific institutions than just some information about the reliability of the SoR given one particular decision to be informed or one question to be answered. The requirements of benevolence and integrity refer to the stable dispositions of the trusted, to their character. If scientific institutions are revealed to systematically work in ways which differ from the expectations of the public, this gives us information about these general dispositions. Publication bias may worry us because it indicates that, generally, the novelty of findings is valued above being sure about the claims under discussion; neglected diseases may cast a bad light on medical research as a discipline because the profitability of drugs seems to be of more importance to the research agenda then the suffering of humans in less well-off parts of the world. If science is revealed to be unproductive or dominated by particular interests although it is sustained by public money and effort, the public may rightly feel betrayed, and not just disappointed. By focusing on discretion, we capture these long-term effects on trust and trustworthiness. Also, the reliability of the object x is connected to worries about the ones that shape it: if we know the competence and dispositions of the producers of some object of trust, we can evaluate the product's reliability as well. If we trust the builders and maintainers of the bridge, we have cause to rely on it. And while in that case, we may still be worried about external factors such as the weather which also shape the reliability of the bridge, the SoR entirely depends on human action, so all its reliability is rooted in the trustworthiness of the institutions that shape it.

This is not everything that could be said about trusting objects; the language used in connection with trust is complex, and has many nuances that are very difficult to capture. For example, consider again the bridge in the mountains: if someone decided to go over the bridge, but relied on a safety rope for reassurance, could this still be a matter of trust, or would it be mere reliance? One may argue that—because the hiker does not expose themselves to risk to the same degree—, they are not actually trusting the bridge. But still, may they not argue that they do, in fact, assume that the bridge will hold, but just generally adhere to the principle of "better safe than sorry"? In the end, these nuances are not essential to account for imbalance in the SoR. While there may be various cogent concepts of trusting objects in themselves, there are good reasons to prefer a concept of interpersonal trust.

With the SoR as the object of trust, we have settled one relatum in the three-place relation of trustworthiness. In the next subsections, I will explore the nuances that appear when we abandon the very general talk about "the public" and "science" and take a closer look at what and whom these collective nouns actually refer to.

4.2.3 *A*: The Trusting Party—Who is "the Public"?

In the general definition of trustworthiness given in section 4.1 the trusting and trusted party were just individual variables, *A* & *B*. In my discussion of the public epistemic trustworthiness of science, I have very generally substituted *A*—the trusting party—with "the public" and *B*—the trusted party—with "science". But whom, exactly, do these general expressions refer to? As for the first, "the public" was chosen as the subject of the trust relationship because of what spawned the interest in problems with the SoR in the first place: the role of science as a source of information. While imbalance in the SoR can be a problem for the decision making of individuals, I have also emphasized the role of science for policy, and it is here where it becomes apparent that the public, i.e., everyone affected by policy, is relevant, and not just some private interest. But the contraposition of public and private does not quite tell us which public is meant here. If we look into the dictionary, even when only concerned with the public as a noun referring to a multitude of people, we are confronted with a multitude of options:

> The community or people as an organized body, the body politic; the nation, the state; the interest or well-being of the community, the common good. […] The community or people as a whole; the members of the community collectively. […] The human race. […] A section of the community, or of the human race, having a particular interest or connection. […] A collective group regarded as sharing a common cultural, social, or political interest, but who as individuals

> do not necessarily have any contact with one another. (*Public, Adj. and N.* 2007)

For my purposes, I want to point out two meanings of "the public" which are important to give an account which is consistent with the rest of the argumentation in this chapter: the first is the public as members of a specific community, the society in a particular democratic nation state. This understanding is important because only in these terms, the political argument that inspired the justice criterion is conclusive: it assumes that the group which is granting scientific freedoms is a community of voters in some democratic society. To explain some cases of imbalance, however, we need to consider another, broader perspective: the perspective of the public as the entire human race. Neglected diseases, for example, may not be a direct problem for the self-interests of most people in affluent nations. Seeing the phenomenon as unfair depends on taking a perspective from which one can weigh the suffering of people who do not have access to newly developed drugs for the diseases that affect them against the benefits to well-off people created by the medications that actually are developed instead. Kitcher, who has been very outspoken about neglected diseases, also favors a broad conception of the term "society" which underlies his ideal of well-ordered science:

> So far, the ideal is not fully specific, since it refers, vaguely, to the range of points of view present in a society without saying how large or small this society may be. Chapter 2 favors a *broad* conception, one that would require scientific significance to be assessed by considering all the alternative perspectives present in the human population, including those of people yet unborn. [...] Plainly, one could draw boundaries more narrowly. One obvious way to do so is to propose that societies are identified with nation-states [...]. (Kitcher 2011, 116)

Not only political issues—which may be concerned with the public in particular political communities such as nations— but also ethical problems may be at the heart of imbalance in the SoR. Ethical argumentation is rarely confined by national borders, and thus, the broad conception of the public as all of humanity will be the second important perspective within this dissertation besides the public as the citizens of a particular society or a national state. But from time to time I will also have to focus on specific parts of the public. The examples of imbalance under examination sometimes concern conflicts between various groups of people who are affected by the SoR; one of the most common contrasts will be those with commercial interests in producing and selling medical products versus prospective patients. What makes science trustworthy may differ for particular parts of the public. Productiveness, after all, reflects the needs of members of society, which depends on individual interest. Science may be productive from the perspective of some, and unproductive according to others. This can lead to conflicts between prospective beneficiaries of science, which is one reason why we also need the criterion of epistemic justice that was introduced earlier in this section.[6] In their work on community-based research, Naomi Jordan, Gust, and Scheman (2011) emphasize that the trustworthiness of scientific institutions varies with respect to certain social groups:

> The blame can be spread around widely, but it's important for universities to take responsibility for their failure to consider the possibility that it might not be rational for members of diverse publics to trust academic research.
>
> [...] Differences that mark inequities of power and privilege, such as race or ethnicity, class, gender, or sexual identity, affect not only the psychological likelihood of trust but also its rationality. It is not rational to trust those whom you perceive to have a track record of disrespectfully treating members of a community you identify with, or whose publicly reported views about your community seem to be either lies or stupid mistakes, or who appear to take no interest in what members of your community have to say to them or in the effects that their views about your community have on the people in it. Given the depth and pervasiveness of social, political, and economic inequality in the United States today, it needn't take malevolence or malfeasance for researchers to act in ways that give rise to such perceptions. Ordinary, orthodox scientific method frequently provides sufficient grounds for mistrust, given the gulf that already exists especially between poor, immigrant, and/or racially stigmatized communities and "institutions of higher learning," which, whatever else they

6. However, also individual members of the public may have a genuine interest in science being fair, or adhering to other moral principles. In that case, science can be considered productive for those people even if it does not focus on their self-interest.

> do, serve to train and educate the ruling and managerial elites and to produce knowledge useful to them. (Jordan, Gust, and Scheman 2011, 188)

While the "orthodox scientific method" may be an issue of integrity, clearly also the requirement of benevolence does have an important role to play in such relations between science and specific communities: institutions and individuals involved with science, both in their statements and other actions, may openly display malevolence towards social groups, and thereby gamble away scientific trustworthiness.[7]

The principle of epistemic justice non-withstanding, certain groups may also be justified in expecting to have more influence on science than others. While—as I will argue in more detail in chapter 6—these privileges should be limited, especially when compared to the basic rights of others, those who provide more resources to science may also be right to expect more in terms of benefits. A company selling pharmaceutical products has to have some benefit if it is to invest great sums into drug research—at least under the terms of most current healthcare and science systems—which are, however, critically discussed within philosophy of science (cf. Brown 2017, Reiss 2017).

Last but not least, note that the public epistemic trustworthiness of science also includes scientists as individuals within the trusting party. Here again, it should be obvious that the requirements for trust can vary wildly from other parts of the public: while it is one of the staples of social epistemology to point out that modern science involves a lot of trust between colleagues, with increased expertise the need to trust in others may decrease where scientists can actually verify the work of others.[8] But also the needs and interests in the outcomes of science as well as expectations about how this work is supposed to be done may be very different for scientists themselves. What is only a fringe-interest to many members of society may spark great excitement among colleagues for whom this is closely connected to their life's

7. I will return to the problems arising from historical problems between science and societal groups in the next section, but also with Heidi Grasswick's idea of epistemic trust injustices in subsection 6.2.2

8. While scientists may often be more able to assess scientific claims than other members of the public, this is not always the case—there is a lot of expertise among non-scientists as well, especially in areas of local experience. Also, expertise is not the only requirement for being able to verify the work of scientists: at the very least, one needs to have the time to actually do so, and thus there still is a need for trust even among experts.

work. These personal interests of scientists should not be disregarded entirely when we think about what research should be considered productive. But we should also—considering the criterion of epistemic justice—make sure that the self-realization of researchers does not take precedence over other, perhaps more basic societal goals and projects.

4.2.4 *B*: The Trusted Party—Scientists & "Science" as an Institution

Scientists may be part of the public, but they are certainly also part of the trusted party, that is, of science. But what does "science", as a general term, refer to? Does it just come down to the community of all scientists? How should we think of its competence, benevolence and integrity? The answer to the second question is a clear "no"; far from all people involved in decisions that shape the SoR are scientists: at the very least, private and public funding agencies have a very strong influence over who receives funding, and thus, ultimately, which projects will be pursued, and how research is distributed. And even if the decision-makers are scientists, they are often asked to make those decisions not as research personnel, but as reviewers, publishers, editors, teachers, and students. Having trust in such a diverse group of people interacting in a variety of different roles with different responsibilities goes beyond trust in just a set of people; the trustworthiness of science is about the trustworthiness of an institution. But what, if anything, distinguishes interpersonal trust from trust in institutions? Russel Hardin has written both on trust in government and other institutions (Hardin 1991, 1998). His general account of trustworthiness, however, is incompatible with the idea of trust as competence, benevolence, and integrity I propose in this dissertation: Hardin's theory of trust is based on the idea that it is the self-interest of the trusted party which makes them trustworthy if it aligns with the interests of the trusting party. In certain cases, the trusting party may expect the trusted to be self-interested; surely, however, it is not the only or even a central expectation that can replace benevolence and principles of integrity entirely. However, his focus on self-interest makes for an interesting perspective on trust in institutions:

> In practice most political institutions are staffed by individuals whose motives are heavily if not entirely self-interested. To gain our trust, they will have to work in our interest. We do this in part by making some officials directly

> answerable to the citizens and in part by making other officials answerable to these. Both these controls are likely to be very loose, but the latter sounds especially weak. What we need to complete the picture is a theory of how the general interest can be served by a government of millions of bureaucrats who are fundamentally self-interested, who are motivated not by unusual public spirit, but only by income and career. (Hardin 1991, 202)

While my view on institutions is not entirely as bleak as Hardin's—in scientific institutions you will certainly find people who are at least also motivated by epistemic considerations, standards of their communities, or even their conception of the common good—he has a point in arguing that when we are concerned with trust in institutions, trustworthiness can hardly be just the averaged competence, benevolence and integrity of the individuals within. Even if we suspected that all scientists are only egoistically motivated, this would not immediately mean that we would have to distrust the institution of science: the structure of science can be such that it diminishes the negative effects of pure self-interest in a system of checks and balances and even channel these motivations in an efficient way towards publicly desirable goals. This structure—be it laws, institutional rules, community standards and practices like peer review, or the relationships between different subordinate scientific institutions such as research institutes, universities and funding bodies—can be seen as the analogue to individual principles and dispositions which may or may not be in alignment with the principles of integrity shared or accepted by the public. Also Helen Longino's idea of objectivity through "transformative criticism" (Longino 1990, 76) may be seen as such a structural principle of scientific integrity: She argues that, even if individual scientists and thus the justification of individual results are biased by contextual values, given the ambitious requirements of "Recognized Avenues for Criticism", "Shared Standards", "Community Response", and "Equality of Intellectual Authority", objectivity can still be achieved at the community level.

Imbalances in the SoR are mainly about cases where this system breaks down or is at least deficient, and the examples often come from authors who propose alternative institutional setups or at least incentives for the individuals involved. This hints at a very important conclusion: the trustworthiness of science as an institution cannot be reduced to the aggregated trustworthiness of the individuals involved. This also has important consequences for interpreting imbalances with the SoR: Imbalance is about the overall distribution of research, and not just about individual findings. While also already individual scientific claims depend on more than just individual scientists, the role of the institutional structure is even more important when taking the perspective of the SoR.

Of course, the behavior of individuals is anything but irrelevant to trustworthiness. In my discussion of the concept "public", I have already mentioned that the benevolence of science as perceived by particular social groups may hinge on particular historical episodes. Hardin (1998) comments on the connection of trust, institutions and benevolence:

> Suppose one has no prior experience of institutions but only of individuals, and one now wonders whether a particular institution is trustworthy. [...] If I have a long history of relatively benign and even beneficial dealings with certain organizations, I can plausibly suppose they are trustworthy with respect to relevant matters. Alternatively, if my dealings have been bad, I can meaningfully say those organizations are not trustworthy. But when I have no experience and no reputational evidence from the experience of others in dealing with those organizations, I cannot say very confidently one way or the other whether they are trustworthy. (ibid., 16)

While members of the public can also be alienated by the structure of institutions—especially when they come in the form of bureaucratic rules that seem to impede the individual's interests—it will also often be dealings with specific representatives of the institution that make lasting impressions in terms of good-will or malevolence. Trustworthiness depends on the structure of the institution, but also on the actions and interactions of individuals: in the case where one scientist acts in bad faith against members of the public, the affected groups have good cause to distrust science. However, it should be noted that such distrust, although warranted, may not reflect the actual trustworthiness of the institution. One individual action should not undermine it entirely; other representatives may work to remedy mistakes made previously—for example, by engaging in science-citizen cooperations—and also the way in which scientific institutions deal with members who violate norms of benevolence and integrity can make an important difference to the contrary.

It should also be noted that, just like with “the” public, the global community of science is only the broadest perspective that we can take on the various specific institutions that produce the SoR on a specific topic. A lack of trustworthiness of science concerning a particular SoR does not necessarily imply that all of science generally has the same deficiency—to generalize from the particular, we need to argue that the same lack of competence, benevolence or integrity is at work on a larger scale. And even with one particular SoR in focus, problems with the trustworthiness of one of the institutions or individual involved may diminish overall trustworthiness of the scientific process, but we should be careful to put blame where blame is deserved: in cases of publication bias, for example, the problem may be with the principles of the authors, the review and publication process, or with the funding bodies, and we should be careful when trying to determine whose trustworthiness is actually challenged by the practice. This also raises the question if the discussion in this chapter applies only to public and not to private research. After all, part of my argument for productiveness and justice as principles of scientific integrity rests on the implications of public investment in science, which does not translate to private funding. I have just admitted that private sponsors might have a reason to expect some rewards from research in which they have invested. But while private research may not be *accountable* to the public to the same degree, if the respective institutions do not adhere to, for example, the principles of productiveness and justice, they are still not a *trustworthy* source of information for the public. This raises yet further reason for concern, because while public and private institutions are different sources of research, they are often intertwined, and—as long as intellectual property does not enter the mix—contribute to the same body of knowledge. Therefore, if we are concerned with science as a trustworthy resource for decision-making, we should not only worry about the structure of public but also private research—and the societal institutions which moderate between public and private science.

4.3 Summary

In the preceding sections, I have presented a framework describing the preconditions of public epistemic trust in science:

Def.: Science is worthy of public epistemic trust iff

1. Science can provide information reliable enough to improve the epistemic situation of members of the public in the given context (**competence**).
2. The discretionary powers granted to science as a provider of information are guided by
   - **(a)** good-will towards the public (**benevolence**)
     and
   - **(b)** principles and dispositions which are shared or well-regarded by the public (**integrity**). This includes:
     - **i.** General norms of ethical behavior and good professional conduct
     - **ii.** Social responsibility as justification for scientific freedom, or more specifically:
       - **A. Epistemic productiveness**: Discretion must be guided by the goal of efficiently producing knowledge which is useful for society .
       - **B. Epistemic justice**: Discretion must not be dominated by the influence of particular interest groups, nor should it be to the unilateral benefit of any such group.

It is based on a general definition of trustworthiness as a three-place relation: a trusting party *A* relies on a trusted party *B* for some *x*. In the context of this dissertation, *A* has been instantiated as the public, which sometimes will be interpreted as the members of a particular democratic society and sometimes as all of humanity, with particular focus on the specific societal groups which are affected by the SoR. *B* has been replaced with science as the scientific institutions which shape the SoR, made up by the individuals in these institutions as well as the institutional structure in terms of interrelations and rules. The object of trust, *x*, in our case, is the SoR that science is entrusted with by the public.

As in all trust relationships, the public epistemic trustworthiness of science comes down to three major requirements: the competence of the trusted, their benevolence towards the trusting party, and principles of integrity which guide the discretionary power given to the trusted by the trusting. As for that third requirement, I have put forward two special requirements on scientific integrity which are based on arguments for scientific freedom: epistemic productiveness & epistemic justice. While the arguments suggest that scientific freedom can promote these goals, granting it does not guarantee that they will be achieved in each and every case; autonomy does not, for example, preclude any science-internal bias towards particular interests. Also, individually, these criteria can only ever provide pro tanto reasons for or against the trustworthiness of some aspect of scientific research. To assess concrete problems, the principles must sometimes be weighed against each other and also against more general ethical considerations, for example, when autonomy in designing experiments conflicts with concerns about human or animal welfare.

In light of all the possible threats to trustworthiness stemming from violations of integrity or a lack of competence, it seems hopeless for science to ever be absolutely worthy of the public's trust. But what should be the consequence of that? It surely does not mean that science is disqualified as a source of information for the public. I have quotes Baier saying that the "trusting can be betrayed, or at least let down, and not just disappointed" (Baier 1986, 235). But a lack of trustworthiness does not imply that trust has been betrayed; only if the trusted manipulates the other party into nonetheless trusting them should we consider it a violation of trust. Also, my concept of trustworthiness is not binary. Instead, the factors of competence, benevolence, and integrity determine to what degree we should be able to trust someone.

Therefore, if we detect a deficiency in one of the conditions of trustworthiness we always have at least three options: firstly, we can try to change science to fit the requirements of trust; secondly, we can stop relying on science as a producer of information to the extent that it violates the requirements; or, lastly, we can reduce our trust in science, but keep relying on it. This last option, as per the explanation of reliance given above, requires the public to take a more active control of the decisions where the trustworthiness of science is considered lacking: for example, we might demand that in cases where we suspect problematic cases of undone research, society needs to take a more active role in agenda-setting. All these options, of course, can be combined in various ways—depending on the problem, and how easy it is to remedy or control it.

I have said in section 3.2 that trustworthiness will be my entry wedge into a discussion of philosophically relevant criteria that will allow me to explain what is problematic about examples of imbalance in the SoR. This is not supposed to imply that imbalances are best explained in terms of impediments to trustworthy science. Following my definition, trustworthiness is no basic property, but instead rests on the requirements given in the definition above; therefore, what is wrong in cases of imbalance is also not in itself a lack of trustworthiness, but a lack of competence, benevolence, or integrity on the side of science. The reason to start with trustworthiness, then, is to give structure to the exploration of this normative background; to allow for a cohesive and consistent explanation of the very vague intuition that imbalance is a problem with conflicts of interest, bias, objectivity, or balance itself. In the present chapter, I have emphasized two special principles of scientific integrity: the principle of epistemic productiveness and the principle of epistemic justice. The two principles provide the basic structure for the next two chapters; in each I will expand upon one of them and analyze the respective underlying normative criteria: In chapter 5, I will discuss the criterion of productiveness as the ideal of efficiently producing knowledge which is needed by the members of society; in chapter 6, I will provide an account of epistemic justice, which explains cases of imbalance as problems with distributions of research which can be considered unfair to parts of society.

Chapter 5

Imbalance and Productiveness

In the second part of this dissertation, I explore different normative concepts which are related to problematic imbalances in the SoR. The productiveness of science, I have claimed in the previous chapter, is one of the major reasons to invest trust in scientific institutions and grant them autonomy. I have already said that productiveness is about efficiently producing knowledge which is useful for society, and, in section 2.3 and subsection 4.2.1, provided a model of relevance and interest which connects scientific findings to the needs of the public. But how does this criterion relate to a SoR and its imbalance? When is productiveness diminished in a way that indicates that there is a problem with the SoR? Not everything that could be said about productiveness in general is of interest to this question; asking what makes some form of organizing collective knowledge production more or less efficient than others, after all, would include nearly all epistemic and practical aspects of science and its products. Imbalance was defined as a structural problem with the distribution of research which is constituted by there either being not enough (lacuna) or too much (overabundance) research of a particular kind, and which does not require that any individual scientific finding is problematic (p. 75). That is, problems with productiveness which are of interest to the discussion in this dissertation cannot be about errors in particular results. The main question for this chapter then becomes: Can the productiveness of science be impaired in terms of an unproductive research distribution, even if the individual results produced by science are reliable—and if yes, how?

I will begin my answer to this specific question by taking up an argument which deals with a very similar issue: Daniel Steel (2018) describes a case of "gaming the dose", where it is conceptually unclear if it constitutes a case of (sponsorship) bias because the claims made in the studies under discussion are not technically untrue; however, they might still lead to errors further down the line. He argues that, although these claims might not constitute problematic inferences in themselves, they might still be considered biased in that they are likely to lead to bad inferences made by the audience of the reports published. This is a very interesting point of comparison: Both Steel's analysis and at least some examples of imbalance are about cases where science can be considered epistemically unproductive, although individual findings do not diverge from the truth. In the next section, I give a short explanation of this concept, which Steel calls "misleading claims", and show that, similarly, also imbalance in the SoR can be understood in terms of posing an epistemic risk for subsequent inferences. In the rest of this chapter, I will discuss two important systematic differences between Steel's case and examples of imbalance, before presenting a framework of possible reasons for imbalance as a problem with productiveness in the last section.

5.1 Imbalance & Misleading Claims

If a scientific claim is true, can it still be considered biased? Steel (2018) motivates this theoretical question referring to a medical trial, where there has been some discussion if it is to be considered a case of sponsorship bias or not: "the LUNAR trial funded by AstraZeneca to compare rosuvastatin (sold by AstraZeneca as Crestor®) to atorvastatin (sold by Pfizer as Lipitor®) for lowering low-density lipoprotein (LDL) and raising high-density lipoprotein (HDL) cholesterol levels in subjects with acute coronary syndrome" (ibid., 121). Steel describes this comparative trial as an example of "gaming the dose" (ibid., 119), i.e., as a case where the dosage at which the two drugs where compared was intentionally selected to give the impression that the sponsor's drug is more effective than the competitor, although at other dose levels the latter might be just as effective or even better. The main finding of the LUNAR trial was that rosuvastatin at 40 mg is more effective than atorvastatin administered at 80 mg. However, or so Steel argues, both meta-analyses and the LUNAR trial itself find that, if the two drugs are compared at a dose ratio of 1 to 4 rather than 1 to 2, there is no difference in effectiveness between the two drugs (cf. ibid). If the sponsor knows in advance what this equivalent dose ratio is, they can design trials that are very likely to favor their own drug by selecting favorable doses to be compared.[1] But even if this were the case in the example of the LUNAR trial, would this make the research and its central conclusion biased? Steel reports two different views within the medicine ethics community: According to one camp, there is no bias, because other studies with independent funding agree that rosuvastatin is more effective than atorvastatin at a dose ratio of 40 mg to 80 mg, and therefore the central claim of the LUNAR trial is, in fact, true. According to the other interpretation, however, it *is* a case of sponsorship bias because the trial is systematically more likely to favor the sponsor's interests because of the experimental design that was chosen (cf. ibid., 126-127).

How does Steel resolve this issue? To start with, consider the definition of bias that is referred to both in the discussion and by Steel itself:

> A bias is a systematic error, or deviation from the truth, in results or inferences. (Higgins and Green 2008, 188)

The point of contention in the discussion is whether something can still be considered sponsorship bias even if, such as in the LUNAR trial, the main results do not constitute a "deviation from the truth". Steel argues that it can, because the conclusions of the studies are not the only inferences that might be of relevance:

> The answer to this question, I suggest, ultimately turns on which claims and inferences are relevant when concerns about bias arise. Just the dose–response relationships reported in results sections of published articles? Or should other

1. Steel admits that even in the case of the LUNAR trial, there might have been other reasons for selecting these dose ratios: "For instance, the authors of the LUNAR trial might argue that comparing rosuvastatin at 40 mg/day to atorvastatin at 80 mg/day is fair, because these are the maximum recommended doses of the two drugs, so this comparison is clinically relevant for patients who need very high doses of statins to reduce LDL cholesterol below a safe threshold" (Steel 2018, 128-129). However, Steel suggests that also the target patient group might have been purposely chosen so that the recommended dose fits the sponsor's interests. Also, even if the choice of dosage were to have been well-intended in this particular case, this does not preclude the plausibility of the mechanism described as misleading claims in general.

> claims in the article be included, such as claims in the abstract or conclusion that may be inconsistent with those results? And what about inferences likely to be drawn by those who read the article and to be promoted by sales reps? If likely inferences of those who encounter health research publications and claims they are used to support fall in the scope of "inferences of studies," then misleading claims can count as biased according to the Cochrane definition even when they are true. (Steel 2018, 120)

The idea here is that although the conclusions of the LUNAR trial are not wrong, they might be misleading for both laypeople—that is, prospective patients—and even doctors: while the sponsors and researchers might be aware that the main claim put forward by the trial does not say anything about the absolute effectiveness of any drug but only about the effectiveness relative to a specific dose, others might not have the background knowledge to realize that, and might erroneously conclude that rosuvastatin is generally more effective than its competitor (and perhaps base their subsequent health decisions on this belief, favoring AstroZeneca's financial interests). Misleading claims, then, are a problem for epistemic productiveness, but not because the claims are in themselves false or simply no knowledge has been produced, but because the form and context of the claims lead to an increased risk of bad inferences in the application of the knowledge, creating false beliefs further down the line; the knowledge is not useful—or perhaps even detrimental—to society.

Examples of imbalance could be understood in a similar way as Steel's misleading claims: both concepts refer to cases where there is a problem with productiveness not because the SoR contains any wrong statements—or at least not necessarily—, but because an imbalanced SoR might lead to problematic inferences further down the line. But there are some differences between misleading claims and imbalance in the SoR that need to be addressed.

Firstly, the concept of misleading claims is still about individual claims in themselves and the consequences of publishing them in scientific studies, not about the effects of lacunae and overabundance that only become apparent when we consider the entire SoR. It should be noted, however, that misleading claims in isolation might not always be problematic: As Steel himself explains, whether a claim leads to erroneous inferences down the line may at least in part be dependent on the background knowledge of the agent making the inference (cf. ibid., 137)—which in turn may very well depend on the SoR and may also be negatively influenced by its imbalance. However, other issues that might lead to erroneous influences are also mentioned, such as "the amount of time a person is able to devote to evaluating a research article" (ibid., 137). As the latter cases are clearly not related to the consequences of research distributions, misleading claims, and imbalance in the SoR should be considered independent phenomena. But this is what makes Steel's example an interesting analogy to cases of imbalance: while they might to some extent share the same mechanism, the differences discussed in the following—especially concerning imbalance and miscommunication—show some important general differences between bias attached to individual claims and biased research distributions.

Secondly, Steel uses the idea of misleading claims to explain a very narrow range of cases of sponsorship bias, while imbalance in the SoR may have other sources than the interests of sponsors, and may sometimes be not even intentional: consider the criticisms of the research agenda in genetics which I dubbed "ignored genes" that I discussed in subsection 3.1.1. Here, structural features of academic careers—such as "the time frames associated with academic promotion" were considered to be the source of imbalance, not commercial interests. But also this second point is no reason to discard misleading claims as an analogy. If misleading claims were caused by some other mechanism than the active involvement of the sponsors, they could still count as problematic. What makes them a form of bias is the systematic deviation from the truth in subsequent inferences. That is, as long as there is a mechanism that explains these errors other than "random mistaken inferences and confusions about science" (ibid., 135), the idea behind misleading claims remains the same.[2]

In the following, I will discuss two more substantial differences between misleading claims and imbalance in the SoR:

1. Misleading claims as presented by Steel are problems with the communica-

2. It should be mentioned, however, that Steel also considers the "inferential asymmetries" (Steel 2018, 135) between producers and users of scientific findings to be an integral aspect of misleading claims. This makes them especially problematic if they are connected to sponsorship bias, as it can lead to epistemic injustice between different societal groups. I will return to this point in chapter 6.

tion of results; that is, if communication were improved, the bias could be avoided. While some examples of imbalance in the SoR can be described as communication problems, others cannot be remedied in this way.

2. Steel interprets misleading claims as cases of alethic risk—"the risk of believing something false". Most cases of imbalance in the SoR also include the risk that some inference diverges from the truth. For some of them, however, it is impossible to explain why they constitute violations of productiveness without also considering the relative usefulness of specific truths for the pragmatic goals of inquiry.

5.1.1 Imbalance as Miscommunication

An essential feature of misleading claims as introduced by Steel is that they result from problems concerning how the conclusions of studies are communicated, and, in consequence, interpreted. In the example above, if everyone relying on the conclusions of the LUNAR trial were to understand that the claim about the superior effectiveness of rosuvastatin has only been demonstrated for a specific dose ratio and may not be safely generalized, the inferences made by the audience should not systematically diverge from the truth, and therefore there would be no bias. As the bias could thus be removed if this information could be clearly communicated, misleading claims constitute a problem of miscommunication. Are cases of imbalance in the SoR—insofar as they are problems with subsequent errors in the first place—also to be understood as communication problems? Would such imbalance always be unproblematic if it were to be communicated clearly to the users of scientific information? There are some types of imbalance that, by definition, are communication problems. Consider what I called cases of overemphasis:

> Overemphasis is meant to signify cases where the focus on a particular kind of research is perceived as an indicator of importance of said topic, while there is no actual justification for that belief (p. 69).

In this description, the recipients of the SoR mistakenly take something as an indicator that is not; if this fact could be clearly communicated to them, the focus on a subject might perhaps be still regrettable for other reasons, but not as a case of overemphasis. Recall the example related to the alleged pause in anthropogenic global warming:

> Research on the pause has thus ultimately reaffirmed the overall reliability of climate models for projecting temperature trends. However, by accepting the framing of a recent fluctuation as a pause or hiatus, that research has, ironically and unwittingly, entrenched the notion of a pause (with all the connotations of that term) in the literature as well as in the public's mind. (Lewandowsky, Risbey, and Oreskes 2016, 730)

The amount of research published on the pause, or so we might say in the language of misleading claims, systematically lead to many misinformed conclusions to the effect that there actually was such a pause in global warming, and, consequently "helps maintain the fiction that the science is still too uncertain to form a reliable basis for public policy" (ibid., 731). The authors explicitly say that a different communication strategy or framing might have avoided this bias in public perception:

> If the fluctuation were instead framed as an instance of decadal variation, then scientists would be able to put the pause to misleading contrarian claims that global warming has stopped. (ibid., 731)

While the relationship between miscommunication and imbalance is quite clear in cases of overemphasis, it is much less straightforward with some of the other types of imbalance I have discussed. Consider, again, the phenomenon of publication bias, where negative results are less likely to appear in publications than positive results. For example, imagine a case where policy-makers, clinicians, or patients might be led to overestimate the effectiveness of an intervention because only significant or positive outcomes concerning the drug under consideration have been published (or rather, a relatively large portion of negative reports remain unpublished). This issue could be framed as a problem with communication by saying that the SoR would not lead to errors or false beliefs if all results had been communicated to decision-makers. This, however, would go beyond what is meant by getting rid of the bias by improving communication: Communicating all the results, i.e., making them public actually removes the imbalance in publication. What I mean by "improving communication", however, is just to improve the background knowledge of the recipients of the SoR, for example, by informing them about the existence of an imbalance.

However, publication bias is also an issue with communication in this second sense: we can hope to reduce the negative effects of publication bias by making people who rely on science aware that publication bias exists and consequently try to detect and account for it. For example, Higgins and Green (2008) mention several ways to address reporting biases such as publication bias, most prominently by funnel plot asymmetry analysis (cf. ibid., 310-319), where statistically unlikely distributions of research outcomes are detected. But as the authors admit, not all cases of publication bias can be detected in this way (risk of false negative), funnel plot asymmetry might have different causes (risk of false positive) and even if it is correctly detected, this might help to qualify the findings, but it cannot replace the value of the studies which have gone unpublished.[3]

Similarly, the problem of undone science may be alleviated by communicating that certain kinds of science have been neglected. After all, turning something from an unknown into a known unknown should improve the epistemic situation of decision makers relying on the SoR. However, even more so than in the case of publication bias, this cannot replace the research that has not been done: with publication bias, it can at least be known what direction the unpublished results would have taken, so that if I know that a research area is affected by publication bias, I may be more skeptical about the reliability of positive results. With undone science, in some cases like the idea of unpatentable research, we do not know if any of the alternatives approaches to medical problems would have actually been effective. Think of the analogy to the digital restaurant guide in chapter 3: Knowing that some part of the city does not get any reviews because there is no mobile internet available in this area—and not just because all the popular restaurants are in another quarter—might make me more willing to visit there and try a new place. However, it still doesn't tell me how my

3. It would also seem that from a certain perspective, transparency might not always improve matters: if it is made known that in a certain field of research publication bias runs rampant, even well-justified claims where there actually is no negative evidence that could possibly be reported may be regarded with more suspicion, and thus move the beliefs of the audience further away from the truth.

dinner would turn out in these places, and it might still turn out terrible, requiring a particularly adventurous sort of dining guest. If I knew, on the other hand, that the app allows some restaurant owners to block negative reviews, it would at least give me the information that these places are likely to warrant a lower overall rating than they currently do.

Overall, while it seems that all the problems with imbalance can be mitigated to some extent by making the recipients of the SoR aware of the problem, only some of them are solely about miscommunication. In most examples, the lack or overabundance of some type of research rests on more substantial issues. But also not all inferential errors caused by imbalance are, as Steel's analysis of misleading claims suggests, problems with deviations from the truth.

5.1.2 Imbalance as Alethic Risk

What is a deviation from the truth in the first place? Steel's analysis of misleading claims makes use of the "epistemic risk"-framework introduced by Biddle and Kukla (2017). In that article, the authors try to differentiate different kinds of risk attached to decisions made during the research process, where the discussion about values in science has relied on the term "inductive risk" alone. The inductive risk argument is often traced back to Richard Rudner (1953), who was discussing the risk arising from making certain decisions in inductive inferences: very roughly put, sometimes deciding if one has enough evidence to justify the acceptance of a scientific hypothesis might be epistemically underdetermined, and in these cases—or so Rudner's argument—"the scientist qua scientist makes value judgments" (ibid., 4) when she lets her inferences be influenced by weighing the ethical consequences of accepting or not accepting the hypothesis. In the more recent discussion, the term has become somewhat of a catch basin for risk arising from all kinds of different activities throughout the research process. In the prominent example from Heather Douglas (2000), for instance, various stages of a study in dioxin research are thought to include inductive risk, including the task of determining whether individual liver slices taken from rats should be considered evidence for the presence of cancer (cf. ibid., 569-572). Biddle & Kukla argue that, instead of using the term inductive risk for various research activities, we should reserve it only for the actual inductive inferences in hypothesis acceptance. For the broader idea of risks that arise at different points throughout empirical research and which include the weighing of values they propose the term "phronetic risk" (Biddle and Kukla 2017, 220), with "epistemic risk" referring to the even more expansive idea of "any risk of epistemic error that arises anywhere during knowledge practices" (ibid., 218). But their framework does not only include the origin of such risks, but also the endpoints: with ethical risk, they "mean the risk of harms (in the broadest ethical sense), although in this context, we are particularly interested in risk of harms that arise during epistemic practices" which is distinguished from alethic risk, which is what they "call the risk of having mistaken beliefs" (ibid., 218).

Steel (2018) describes misleading claims as instances of epistemic, and more specifically, phronetic risk: it refers to risks arising from the value-laden deliberations of researchers about the consequences of choosing a particular empirical research design in studies like the LUNAR trial and about how to frame their conclusions. More importantly—for the purposes of this section—, Steel interprets misleading claims in terms of deviations from the truth; his principal worry concerns the decision-makers who depend on the results of medical research who may form mistaken beliefs regarding the effectiveness of medical interventions. As they appear to be about the risk of having mistaken beliefs, Steel considers misleading claims a source of alethic risk.

Before I apply these ideas to imbalance in the SoR, however, some more theoretical observations are in order. Firstly, while alethic risk—that is, truth-related risk—was so far described as being about mistaken beliefs, this is not the only way to conceive of a risk that is about deviations from true belief. In his veritistic framework, Alvin Goldman (2003) distinguishes three broad epistemic states that are of note: "The first state constitutes *knowledge*, the second *error*, and the third *ignorance*" (ibid., 89), where this constitutes a continuum from being absolutely right about something to being absolutely wrong. Whenever we move away from the state of knowledge, this might be considered alethically problematic, even if we are only in doubt, but have not yet committed to an erroneous belief. In the example of the alleged pause in global warming given above, we already encountered both notions: the overemphasis of this topic was thought to have reaffirmed mistaken belief—i.e. error—in this pause, but also, more generally created doubt about the truth of the claims made by climate science—which comes down to ignorance as discussed by Goldman. In contrast to Goldman, however, I would not call the epistemic state which is between true belief and error ignorance, but uncertainty.[4]

"Ignorance" I reserve for unknown unknowns, i.e., for cases where we do not even know that there is something that we may want to know. The phenomenon of unpatentable research may motivate this as a possible reason for imbalance: if patients and doctors do not read about alternatives to drug-based medical interventions, they may not even consider them a possibility. I suggest, therefore, that we consider a decision in the context of science alethically risky if it threatens to lead a) to false belief, b) to uncertainty or doubt about a true belief, or c) to ignorance in the sense of being unaware about the existence of a relevant question or decision-option.

One may wonder why option c) above mentions decision options besides relevant questions. This is because, as I will argue below, to account for the different problems with the productiveness of science that are caused by imbalance in the SoR, we need concepts beyond alethic risk. As explained above, Biddle and Kukla (2017) offer ethical risk as an alternative consequence-related concept connected to epistemic risk. Their definition of ethical risk seems insufficient both on theoretical grounds—avoidance of harm is a rather limited understanding of ethical consequence—and also in how it appears in their discussion: in their visualization of the "geography of epistemic risk" (cf. Figure 5.1) only a small fraction of inductive risk is accounted for by ethical risk, without any argument for in what other way inductive risk may be problematic. The authors do admit that the diagram "is intended to be illustrative and not exhaustive" (ibid., 221). However, they do not even commit to answers about relatively elementary questions, such as "should alethic risk overlap with ethical risk" (ibid., 221).[5] The concept of ethical risk provided by Biddle and Kukla thus does not appear to be a very promising criterion for the analysis of imbalance.

Moreover, the two concepts of alethic and ethical risks are neither clearly dis-

4. I discuss this difference to Goldman in more detail in subsection 5.2.1

5. I would argue that if we do not want to claim that being mistaken or ignorant about some truth can never have any negative ethical consequences—which appears to be quite absurd—they should, indeed, overlap. This will be demonstrated in the discussion of the examples below.

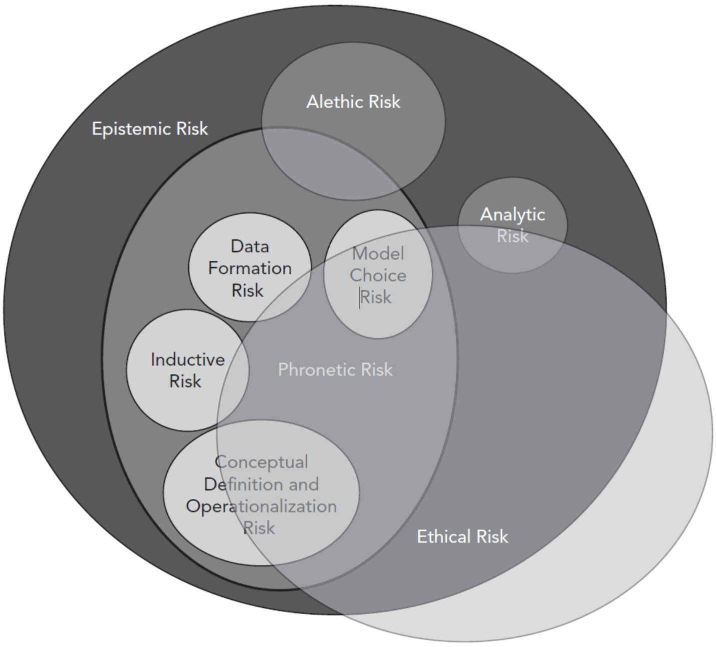

Figure 5.1: "The geography of epistemic risk", reprinted from Biddle and Kukla (2017, 222).

tinct nor would a framework based on them be comprehensive. If knowledge has intrinsic worth, then alethic risk is ethically worrisome. But also, not all concerns beyond worries about true belief are truly ethical issues—they may also be just about instrumental rationality. Think back to the account of interests and relevance in subsection 4.2.1. There, the main example of interesting or relevant research questions was connected to an action-guiding argument: a decision about banning a certain drug or approving it which concludes in a recommendation for action (cf. Figure 4.1). Evaluating the outcome of this inference could be considered an issue of ethics—i.e., we could ask if taking one action or the other is morally right. However, we can also simply ask if it is hypothetically right; that is, we may ask if it is right given a specific goal or interest. Already in the case of drug approval, some may object that seeking to avoid risks to the public in the approval procedure is not necessarily the only morally relevant goal; once we have set this goal, however, we can ask if the SoR contributes to this goal or impedes it. But even in cases of pure self-interest, we can distinguish research which is helpful to these interests, and research which is not. We can ask, for example, what research questions would be the most relevant or interesting for a company that produces these drugs, and is mainly interested in generating the most profits. Beyond alethic risk, we can thus be worried about risks to productiveness both in terms of truly ethical considerations or about negative consequences to any interest a member of the public might have. For now, I will treat all such cases of risks to successful decision-making—again, in terms of error or undue withholding of judgment—the same. I propose to call such risks practical risk. I will return to the distinction between the alethic, the ethical, and the *merely* practical at the end of this chapter. In the remainder of this section, I argue that cases of imbalance in the SoR can 1) sometimes be explained in terms of alethic risk alone, but also 2) sometimes have to be understood as a mix of alethic and practical risk, and 3) may even sometimes be explained solely in terms of practical risk.

As for the first category, imagine, once again, a case of publication bias, where for some psychological effect under investigation, there were both studies that confirmed its effectiveness, but also studies that did not find a significant effect. However, because the authors of the latter studies did not consider it worth their while to try and get them published—perhaps because they assume that they would not be taken up by any prestigious journal or because they do not think it will further their career—only the studies with positive results get published. Why might that be problematic? Let us assume that in this case, the studies that were published are actually methodologically sound, i.e., among other things, the experimental design was well done, the statistical analysis relied only on accepted tools, and the authors did not try to put a particular spin on the results. However, as the other results are not published, the overall evidence synthesis—perhaps in a meta-analysis of all published studies—may still be skewed towards the positive. The audience of the SoR—that is, other scientists and interested parties, but in this case, also the researchers themselves—may be misled in their beliefs about the actual existence of the psychological effect, that is, they may believe it exists when it actually does not, or at least be less skeptical about it than they should be. It should be uncontroversial that in such a case, publication bias poses an alethic risk in terms of mistaken beliefs or ignorance in terms of uncertainty. We can, of course, also consider the ethical or practical risk posed by inaccessible research.

As Chan et al. (2014) warn:

> Selective reporting of studies means that fully informed decisions cannot be made about care of patients, resource allocation, prioritisation of research questions, and study design. This ignorance can lead to the use of ineffective or harmful interventions and to wasting of scarce health-care resources [...]. (ibid., 3)

Publication bias, in many cases, has serious consequences that are not intrinsically about true beliefs or ignorance of questions or options, and one may argue that these are the more worrying effects. However, even if such biases only were to lead to errors concerning questions of interest that people are simply curious about, we can describe them as a problem, just relying on a concept of alethic risk.

Consider, in contrast, the example of neglected diseases. Philip Kitcher (2011) describes it as follows:

> Consider contemporary biomedical research. Most of it is carried out in affluent societies, and almost all of it concentrates on diseases afflicting people in those societies. [...] Contrast the distribution of disease research with the statistical data in worldwide disease and disability. Diseases that cause a vast amount of human suffering, particularly among children, receive only a tiny part of the investigative effort. (ibid., 121)

Certainly, also this example implies that the recipients of the SoR may be led to erroneous beliefs or ignorance about something: if some disease is not in the focus of research, our understanding of it will be lacking. But unlike in the example of publication bias, here it is not immediately clear that the overall alethic risk would be reduced by an alternative approach. In cases of publication bias, if studies of high quality concluding in negative results would be published, subsequent inferences about, for example, the existence of some effect under investigation would be improved. In the case of neglected diseases, the idea is not to just increase the overall level of funding, so both the hitherto neglected diseases and the diseases currently in focus could be researched; rather, or so the authors, we should redistribute funding from the diseases afflicting people in affluent societies towards the neglected ones. While this would in all likelihood decrease alethic risk concerning the latter, it would also increase alethic risk concerning the former. If we still believe neglected diseases to be a problem, we cannot explain that intuition with the criterion of alethic error alone. What makes the important difference between the possible research distributions are not directly the consequences for true belief, but the effects on practical risk:

> [A]t least insofar as disease problems are seen as comparably tractable, the proportions of global resources assigned to different diseases should agree with the ratios of human suffering associated with those diseases (Flory and Kitcher 2004). Thus if the disease burden associated with a form of respiratory infection is twice that of a specific type of cancer, and if there are approaches to both diseases that are roughly equally promising, then the funds assigned to the respiratory infection should be approximately twice those given to the cancer. (Reiss and Kitcher 2009, 263)

Even if both lines of research are "roughly equally promising"—that is, if they are roughly equally likely to produce results that could improve our knowledge and thus reduce alethic risk—, the research on the respiratory infection is preferred. This is because of the amount of suffering that could be avoided, that is, because it contributes more to practical decisions about health than the alternative cancer research. Without this difference in suffering avoided; that is, without a difference in practical risk, it would not be clear why this case constitutes a problematic imbalance. Of course, in this case, the difference in practical risk depends on us lacking certain knowledge, which then leads to an increase in practical risk. As I have claimed before, we should generally only be concerned with relevant, i.e., significant or interesting truths, and not with just any knowledge. From this perspective, we could even say that focusing on the respiratory disease is alethically less risky, because we reduce the risk of erroneous inferences or ignorance about *more relevant* questions. The point here is, however, that this difference in relevance only appears when we consider the level of practical decisions, and thus the criterion of practical risk is indispensable in explaining why the example of neglected diseases constitutes a problem with productiveness[6].

6. Neglected diseases, however, are not only described in terms of the amount of suffering caused—

A third variant in which alethic and practical risk are related to productiveness might be most controversial: Can there really be cases in which an imbalance in the SoR may pose a practical risk without this negative influence on decision-making being rooted in false belief? At least beyond the scientific realm itself, evidence is not always only used to actually justify belief, but can also be a formal requirement for taking some action. Scientific studies sometimes are part of expert assessments, where the results have a direct effect on decision-making. If this decision procedure is highly formalized, evidence in favor of a decision option or against it may be enough to lead to a decision, without anyone actually having to wrongly believe that the decision is based on facts about the matter. Think of drug-approval procedures: when the FDA decides about accepting a new drug, "the FDA requires 'substantial evidence' of drug safety and efficacy, and interprets this as needing at least 2 adequate and well-controlled Phase III trials with convincing evidence of effectiveness" (Van Norman 2016, 178). Given this action-guiding framework of drug acceptance and mechanisms such as publication bias and suppression of evidence—which might lead to studies contesting effectiveness or reporting side-effects being unavailable—the FDA might make an error concerning the acceptance of a drug, where error is supposed to mean that it fails to bar an ineffective or unsafe drug from the market—which could be understood as the pragmatic goal of this process. If we were to understand this as an algorithmic decision principle—"If there are two studies confirming effectiveness and no other studies reporting side-effects, then approve the drug"—nobody involved at the FDA actually needs to hold a false belief about the drug for it to be erroneously accepted. Of course, one may object, this is a grossly simplified picture, and if there was cause for doubting the efficiency or safety of some drug, respective expert opinions may still prevent practical errors: "FDA reviewers will evaluate clinical data, analyze drug samples, inspect the production facilities, and check proposed labeling. [...] The FDA often convenes advisory panels of experts to review the data, and usually follows panel recommendations." There is in fact room for doubt to influence the

which hints at a problem with practical risk and productiveness—but the example also suggests an inequality concerning *who* suffers from the distribution of research; that is, it also conflicts with a principle of epistemic justice. This twofold wrong of neglected diseases will be a central issue in section 6.1.

outcome. Still, I would suggest that in this case, the formal requirement in terms of the number of required trials puts some direct pressure on the decision-making process and thus pose a practical risk, which, to some extent, does not depend on the beliefs of the agents involved. This is important because it should make us wary of highly formalized decision-processes depending on scientific data. They are vulnerable to imbalance in a different way, especially were we might suspect active manipulation.

Still, one might be worried that it may either be not necessary, not helpful, or perhaps even impossible to distinguish between alethic and practical risks. Interestingly enough, this worry goes in both directions: on the one hand, some may argue that productiveness is always about practical risk; on the other, some may claim that all cases presented in this chapter refer to false belief. As for the first alternative, we might say that all the inferences thought to be problematic are about what to do; it is only that what I described in terms of alethic risk refers to actions and of a particular kind: accepting some proposition as true, rejecting it as false, or withholding judgment. After all, I myself have suggested that both scientific inferences and the action-guiding framework can be reconstructed in terms of arguments, that is, as forms of judgment. But there is a relevant difference between coming to believe in a proposition and explicitly making some judgment about what to believe. This distinction has been upheld in the discussion about values in science. Consider this example of an aspect of scientific practice which may constitute epistemic risk:

> So, for instance, as we develop perceptual skills, we see and classify what we see in distinctive ways; a radiologist does not see an MRI or ultrasound reading the same way a layperson does. But when she sees an abnormal growth or whatever it may be, her vision already encodes a balancing of values; if her perception is extra- sensitive to abnormalities, it will catch more false positives and fewer false negatives, and vice versa. Her visual examination is not plausibly an inductive inference from statistical data. But values are built into her perceptual episodes during the course of her epistemic practice. (Biddle and Kukla 2017, 221)

What is made explicit in this quote is the fact that epistemic activities which philosophers reconstruct as judgments are sometimes just that: rational reconstructions of complicated psychological processes. These reconstructions may then be used as models that can tell us what the reasons for believing some thing or the other *should be*. In reality, however, coming to believe in some proposition may not always be adequately described as an activity, but rather as a largely passive process. Still, we may wonder if this subtle difference is of any consequence. Consider again the difference between coming to belief a thing and the three kinds of actions mentioned above: accepting or rejecting a claim and withholding judgment. Especially concerning the role of information provided by science and decision-making—which is a major theme of this dissertation—it should be clear that the distinction leads to interesting questions: Given how certain a scientist is about a specific proposition, what claims should she publicly endorse? This difficulty is at the heart of many discussions about values in science and inductive risk, and the distinction disappears if we claim that all the negative aspects of imbalance come down to just what actions to take.

The second alternative, then, would be to argue that all problems with productiveness can be explained in terms of false beliefs. While alethic risk, as I described it, is about the truth of descriptive claims, what I call practical risk would be about normative truth: practical risk, then, would be the risk of having false beliefs about what actions to take. The reason why I do not describe the difference in these terms is mainly that I do not want to commit myself to an ontology which implies the existence of such things as normative truths. Also, this metaphysical worry taken aside, it seems that proponents of such a position might still accept that there is a difference to be made between the alethic and the practical, just that the demarcation would not be drawn along the lines of true belief versus right actions, but between different kinds of beliefs. I do not believe that anyone would deny that there is an interesting distinction to be made between asking—for reasons of curiosity—what killed the dinosaurs and asking—as a doctor or patient directly affected by the answer—if some drug should be approved or not. While the first is just problematic in terms of any intrinsic value we assign to being right about our beliefs about the universe, the second is about consequences of relying on the SoR which go beyond intrinsically truth-related concerns.[7]

7. Note that the difference between the alethic and the practical is not the same as the distinction between epistemic and non-epistemic values, which has been criticized in connection with the values in science debate (cf. Rooney 1992). The distinction of alethic versus practical risk is about true beliefs

5.2 Towards the Framework

In section 5.1, I have introduced the idea that many of the problems related to productiveness that are caused by imbalance in the SoR are about both alethic and practical errors in the inferences made by the recipients of the SoR. I have then compared examples of imbalance to the concept of misleading claims, finding both similarities and differences to individual cases. In this section, I will reconstruct a framework that allows us to use the criterion of productiveness to explain why specific instances of imbalance are problematic. Before I can make this account more precise, however, I will reply to some difficulties in applying what I have presented so far: first, I will discuss what type of framework I plan to provide, second, I will counter the argument that cases of alethic risk might reduce to problems with individual scientific findings, and third, I will discuss the possible objection that the concept of diversity has been neglected in my discussion of productiveness.

5.2.1 Challenges

The type of framework. What kind of framework is needed to account for imbalances in the SoR? Evaluating the productiveness of distributions of research should be comparative rather than attached to a particular distribution in isolation. That is, we cannot call a SoR imbalanced just because there is some alethic or practical risk; even if we restrict ourselves to certain questions of interest, there will always remain the chance that *someone* will make problematic inferences based on the available research, both because science will never provide us with perfect knowledge about a subject matter and because making the right inferences does not only depend on the epistemic resources provided by the SoR. Furthermore, we do not live in a perfect world in the first place, in the sense that the resources for research are limited, and we may often not be in a situation where, for example, a lacuna can be filled just by doing more research without neglecting other questions. Therefore, to be able to claim that some case constitutes problematic imbalance in the SoR, we have to argue that due to the current research distribution having different features in terms of communicability or

versus right actions, while the epistemic/non-epistemic distinction is about values that promote the attainment of truth in contrast to all other values—often social or ethical ones.

ignorance than another, the former poses more alethic or practical risk than possible alternatives.

But what is a "framework" in the first place? One might imagine a set of formalizations, that allows us to determine quantitatively how productive some SoR is. Goldman's veritistic framework—which we first encountered in subsection 4.2.1—is an example of such an attempt. Social practices are compared in their relative ability or propensity to increase the overall V-value: Goldman asks whether by engaging in some practice rather than another—for example, one way of organizing science compared with a second—"the *aggregate* level of knowledge of an entire community" (Goldman 2003, 93) is increased or decreased. The level of knowledge is calculated by considering the true answers to the questions of interest and asking, for each member of the community, what their actual degrees of belief in these propositions are like: if p is true, and DB(p) = 0.00, there is no veritistic value, if DB(p) = 0.50 the value is 0.50, if DB(p) = 1.00 the value is 1.00 as well, and so on. To arrive at the overall level of knowledge, we are to aggregate the value for the individual questions and agents, for example, by forming an average, although Goldman suggests we might also consider alternative measures (cf. ibid., 93-94). There are some obvious problems in applying this neat and seemingly precise framework to the ideas above: I am not only concerned with knowledge, or rather, true belief versus false belief, but with other risks as well. Already concerning alethic risk, this creates problems: as David Coady (2010) claims, while—in my terminology—alethic error and uncertainty may be put on the same scale, the question of ignorance cannot be captured by the same measure. Practical risk adds yet another dimension to the problem. It is already unclear how only the value of true beliefs and ignorance could be combined into one quantitative measure; I am very pessimistic concerning finding measurements for each of the above risks that, in the end, would be commensurable.

But there are also intrinsic problems with Goldman's account. Even when restricting ourselves to the first two types of alethic risk, there are many questions about how to determine veritistic value: How do we measure degrees of belief? This question also haunts proponents of Bayesianism, and it does not appear that this debate has arrived at a satisfactory conclusion.[8] How do we attribute changes in belief to competing social practices—or, in our case, distributions of research—, especially as long as alternative candidates remain hypothetical and thus untested? Goldman does touch upon the subject of attribution himself (cf. Goldman 2003), but he only discusses the question of how to deal with multiple concurrent social practices. He does not tell us how to find out if a particular degree of belief was due to the distribution of research, some other feature of science, something about the subject's psychology, or a myriad of other possible influences. Lastly, how do we weigh differences in interest or relevance between alternative questions or research projects? Goldman's veritistic value only concerns interesting truths, and not just any truth. I, so far, have used the concept of interest as if it were binary: either something is actually, hypothetically or instrumentally interesting, or not. While Goldman does recognize the importance of what could be called "degrees of interest" in addition to degrees of belief, he also has two worries about including them in his framework. First, he is concerned that focusing on the amount of interest a question arouses may obscure the difficulty of and thus the "intellectual skill" needed for answering the question—a feature of veritistic value that seems not particularly relevant to imbalance in the SoR. Second, he is worried that considering the degree to which people are interested in certain questions "may reflect factors that do not properly belong in an epistemological analysis" (ibid., 95). That is, he fears that veritistic value—or in my terms, alethic risk—may be mixed up with other, practical considerations. But even if we believe that there are purely epistemic interests, it should be apparent that different degrees of interest matter: Consider again the notion of instrumental risk, where the actual interest in one question implies an interest in other questions relevant to that actual, primary interest. For example, a person may ask themselves

8. One of the standard ways of determining the degrees of belief of an agent is via their betting behavior, i.e., by their willingness to accept a bet that the proposition under consideration is true, or their judgment what betting ratio would be fair concerning the proposition. Cf. Huber (2008) for a discussion of various arguments against the idea that betting behaviors are a convincing proxy for degrees of belief, e.g. because this account ignores the issue of risk aversion (cf. ibid., 4-5) or more formal problems such as the lottery paradox (cf. ibid., 9-10). This is not to say that there is no way to measure degrees of belief in principle; rather, there is an ongoing discussion with competing theories of belief, and a unilaterally convincing account has yet to be defended.

why the dinosaurs became extinct. This main question may imply some instrumental interest concerning subsidiary questions, such as "Was there a cataclysmic event in the past such as a meteorite strike at the time the great reptiles disappeared from the fossil record?", but also, for a specific fossil "How did this particular dinosaur die?". Arguably, both subsidiary questions could be interesting to our main inquiry, but it could certainly be argued that the first may contribute more to our overall inquiry and is thus more relevant. Different answers contribute to the primary goal to varying degrees, making them more or less interesting. Even more to the point, answering the primary question seems to be of more veritistic value than answering any of the subsidiary ones—they are only of value in terms of contributing to the overarching goal, after all. Goldman admits to this specific difficulty in a footnote (cf. Goldman 2003, 99-100) and also, despite his reservations, accepts the general importance of accounting for different degrees of interest:

> A social practice that systematically delivers information on topics of mild interest to an agent while regularly concealing or masking evidence on topics of core interest is an epistemically unsatisfactory practice. (ibid., 95)

However, this point is not adequately reflected in his veritistic framework:

> Intuitively, more V-credit should be given for true answers to the primary question than for true answers to the subsidiary questions. It is not obvious, however, exactly how to quantify these matters. In general, I have not developed a full-fledged "calculus" of V-value here. What has been developed, however, should suffice for purposes of the book. Further refinements might be added in the future. (ibid., 100)

While we may or may not grant this simplification when only concerned with truth, in the case of productiveness as a criterion for assessing research distributions it would be a fatal flaw to disregard different degrees of interest: when we ask if a SoR is conductive to efficiently producing knowledge which is useful to society, we need to consider the relative instrumental usefulness of research invested in one rather than another question or producing one rather than another kind of result.

Overall it seems that, while a precise, quantified notion of productiveness would be helpful in comparing different SoRs and the distribution of research within them, for all the reasons outlined above, a framework as presented by Goldman—while it may serve to clarify particular aspects—is too simplistic to account for all the relevant aspects. I also assume that "further refinements" will not, in the end, be able to resolve these problems, because some worries—such as the incommensurability between different kinds of risk—appear to be fundamental, and not just a matter of precision or complexity of the framework. What, then, can be expected of the productiveness-framework presented in this chapter? Rather than thinking of assessing the relative productiveness of two SoRs in terms of computing and comparing one unified measure, I suggest we conceive of this task as an issue of complex argumentation, where different reasons for preferring one distribution over the other must be given, and in the end weighed by the people affected according to the weight they assign to these arguments. The aspects of productiveness discussed provide an overview of different reasons that may be considered when making such a judgment. Also, the conceptual framework provided is not supposed to provide an instruction manual for assessing and choosing research agendas; the question it was supposed to help answer is why we intuitively consider certain cases and phenomena to be problematic imbalances in the SoR. If the framework presented is to be convincing, it needs to provide a comprehensive list of the possible reasons we might have to criticize a SoR's productiveness that can account for all the examples under consideration.

Productiveness, imbalance, and individual results. In section 3.2, one of the requirements I gave for some mechanism to count as a source of imbalance in the SoR was that it "does not require that any individual scientific finding is problematic" (75). One may be worried, however, that at least some of the examples mentioned in this chapter reduce to problems about the quality of individual claims; that is, if the criterion of productiveness is explained in terms of these cases, it might be too broad to address imbalance as a problem in its own right. On a general level, one may be concerned about the analogy to misleading claims: after all, the problem identified by Steel again is about bias in individual inferences. His case, after all, is not based on problematic distributions at all. But in the majority of cases discussed—and in fact, in Steel's examples as well—not individual *scientific* claims were in focus, but the subsequent inferences made by the audience. Still, at least in some cases, the notion of "audience" might also refer to other scientists, and the biasing influence of the SoR may be about the truth or reliability of individual scientific claims. Think, once again, of publication bias: usually, this phenomenon is discussed in connection to scientific judgments about the justification of particular scientific hypotheses, for example about meta-analyses, which aggregate the findings concerning some effect under investigation.[9] Is this not a case where imbalance is reduced to a problem with individual scientific results?

This interpretation of publication bias, however, is somewhat misleading. Imagine a situation where there is publication bias concerning the effectiveness of some medical intervention. However, nobody has attempted to make any claim about the overall effectiveness. Has the problem with publication bias now disappeared or, perhaps, not yet arisen? The answer should be no; the problem with publication bias is not inherently with a hypothesis being formed based on the SoR, but with the distribution of research itself. We can also think about this in interventionist terms: Would we be more likely to be successful in removing the problem with publication bias by changing a particular inference, or would it be more helpful if the SoR would change so that both the negative and positive results were available? While worries about publication bias might often be about a concrete scientific hypothesis which is skewed towards the positive, there is never only one specific inference which is at risk. Even without an unreliable meta-analysis as an intermediary, patients and doctors may themselves form various opinions based on a SoR skewed by publication bias, and it is this general risk attached to a particular distribution of publications that makes it worrisome. It simply is a deficient source of information, no matter the particular outcome under consideration. Therefore, we can only deal with the problem entirely if we balance the distribution, and allow for the negative outcomes to be published.

9. It may also be questioned if the individual results the meta-analysis aggregates can count as unproblematic if the inclusion of all research, published or not, would change the overall assessment. After all, the negative results would, to some extent, contradict the findings of the studies that have actually been published, so one could argue that the published findings are problematic because they overestimate some effect—and thus deviate from the truth. However, as long as the studies fulfill the respective methodological standards, they should not be criticized for such deviations; or at least not as a problem with publication bias, but rather perhaps as a general criticism of statistical methods.

This answer can be generalized: for something to count a case of imbalance, the problem to be resolved needs to originate with the overall distribution of research, irrespective of the existence of a particular scientific finding that deviates from the truth. However, an increase in the risk of such claims being made may very well be one reason to argue that a particular distribution of research is problematic.

Before I review what these considerations imply for the framework, there remains one more worry to be addressed. In the list of values discussed above, we might feel that another is missing, which was described in chapter 4 as one of the major reasons why scientific freedom can be seen as a way to increase productiveness: diversity.

Productiveness and diversity. The definition of imbalance in the SoR in terms of lacunae and overabundance seems to suggest diversity as the core explanatory concept: Imbalance is either about a lack of particular types of research or about focusing too much on a particular type of research. In both cases, we may criticize the SoR for not being diverse enough. But what is explained by understanding imbalance as a problem with diversity? Why should we care if something is diverse or not? And what does it have to do with productiveness? The last question hints at a vital distinction: diversity may be both intrinsically and instrumentally important. The first way of thinking is connected to what will be discussed in the next chapter: diversity may be a measure of representing the various view-points in a heterogeneous society, and may thus a matter of intrinsic participatory justice. As a value connected to productiveness, however, diversity appears as a purely instrumental value. Different arguments have been advanced in favor of diversity: As Kitcher (1993, 2011) most prominently argues, methodological diversity can make it more likely that the scientific community will achieve the goals of inquiry. This is justified by the diminishing returns of piling all resources—that is, for Kitcher, mostly researchers and their time—on the same method:

> [I]f there are diminishing returns to additional investment in any particular strategy; that is, if adding one more scientist to the pursuit of that strategy raises the probability of the strategy only slightly, and if the probability of a different strategy's being successful being successful, given pursuit by a single scientist, would exceed that slight amount, it is better to divide the labor. (Kitcher 2011,

194)

Kitcher's argument above can not possibly apply to all cases of imbalance directly: while some of the categories of research types discussed in section 3.1 can be interpreted in terms of methods or research strategy—think of the methods favored by the EBM movement—, others cannot. For example, cases like unpatentable research or ignored genes are not about which methods, but about what topics are neglected. This does not mean, however, that a version of Kitcher's argument may not also be put forward in those cases: as long as there is a common goal to be served by inquiring into different topics—for example, if drugs and alternative methods or different kinds of proteins are studied to provide a remedy for a specific disease—it is still plausible that at some point the returns of investing in only one of these lines of inquiry will decrease. But still, with some instances of imbalance—e.g. about which actors can contribute to the research—it is at least less obvious how this can be a problem in terms of diminishing returns.

But there is another line of argument, which can endorse a principle of diversity as instrumentally beneficial for epistemic productiveness. Consider this statement by James Robert Brown concerning his example of unpatentable research:

> [...] [I]t is the job of philosophy of science to make the methodological point that without seriously funded rival approaches, we will never know how good or bad particular patentable solutions really are. The epistemic point is commonplace among philosophers. Evaluation is a *comparative* process. The different background assumptions of rival theories lead us to see the world in different ways. Rival research programs can be compared in terms of their relative success in the long run. But to do this, we need strong rivals for the purposes of comparison. (Brown 2008, 199)

This "epistemic point" is made more explicitly in Brown (2001):

> No longer do we think that theories can be tested solely by the evidence. Rather, theories can only be evaluated with respect to their rivals. Given some body of evidence, we can say that T_1 is a better theory than T_2; but we cannot say that T_1 is true unless T_1 is chosen from a more or less exhaustive pool of candidates. (ibid., 185)

Brown is not alone with this position in philosophy of science. Lakatos, for example, argues that we cannot have objective judgments of individual scientific theories based on empirical evidence alone. According to him, scientific theories are not only fallible but also—chiefly because there is always another possible way to explain any experimental result—empirical evidence can never directly falsify a theory: "Thus *we cannot prove theories and we cannot disprove them either.*" (Lakatos 1976a, 16) But also other criteria of individual theory evaluation are not enough: "Neither the logician's proof of inconsistency nor the experimental scientist's verdict of anomaly can defeat a research programme in one blow" (Lakatos 1976b, 113). Consequently, Lakatos claims that we can only rationally discard a theory or "research programme" after long-term comparison with a better one:

> Can there be any objective (as opposed to socio-psychological) reason to reject a programme, that is, to eliminate its hard core and its programme for constructing protective belts? Our answer, in outline, is that such an objective reason is provided by a rival research programme which explains the previous success of its rival and supersedes it by a further display of *heuristic power*. (Lakatos 1976a, 69)

But how does this epistemic argument about scientific theories and methodologies relate to the criterion of productiveness? I am not directly concerned with theory evaluation in this dissertation, after all. Certainly, we can admit that lacking alternatives for comparison weakens the justification of our beliefs and thus increases alethic risk. A similar point has been made with what Anke Bueter (2015) calls "value-laden blind spots in the scientific community" (ibid., 18):

> Non-cognitive values can affect which data are given and which theories are pursued. Hence, they have an impact on theory evaluation via the questions, which data a theory needs to account for, and against which theoretical alternatives it has to excel. Even if a theory is currently empirically adequate and is the best alternative in light of its rivals, it may still be the case that it would not be accepted if there were other data or other rivalling theories. These non-existent data and rivals might be non-existent because of values in discovery making other questions or aspects seem insignificant or even invisible. (ibid., 21)

This account is very close to an idea of imbalance in the SoR: here, a lacuna in terms of available theories or pursued questions—systematically caused by the prevalence of certain non-cognitive values—may lead us to misjudge the theories or hypotheses which are available, leading to alethic and consequently also practical errors. But especially when we have more practical goals in sight—such as, in the case of unpatentable research, dealing with diseases—can we not also be satisfied with evaluating a scientific finding in terms of success? Kitcher, in connection to this definition of significance in terms of problems worth pursuing, explains:

> Those problems are *adequately solved* when an item is produced that is close enough to the type sought to serve the purposes that confer significance to the problem. If the problem is to answer a question, an adequate solution is a statement "true enough" to enable those who have it to achieve whatever ends made the question significant. If the problem is to produce a new vaccine, an adequate solution is one providing acceptable protection against the pertinent disease. (Kitcher 2011, 105)

If we connect the success of science to a particular need—as the criterion of productiveness generally does—one may argue that, questions of pure curiosity aside, the epistemic point above becomes moot. However, this ignores the fact that, for most significant issues, success is not binary: We can be more or less successful in satisfying the needs of society. While a vaccine may provide "acceptable protection against the pertinent disease", it can have higher or lower success rates, it may or may not cause side-effects, and may be more or less affordable, only to name a few criteria. How good a solution really is may also only become apparent when we compare it to the possible alternatives. Also, comparisons with other alternatives may reveal criteria which have hitherto gone unnoticed:

> Not only are theories evaluated by means of the evidence relative to their rivals, but what counts as evidence may depend heavily on what rival theories are being considered. (Brown 2001, 185)

To return to the issues of unpatentable research, once we seriously consider alternative approaches to drug-based medicine, we may also ask different questions about the solutions: when we consider alternatives like sports, diets or social aspects of disease, we may evaluate the provided answers not only in terms of their ability to cure the disease, but also to prevent it altogether, or, for example, in terms of the impacts on social relationships in the patients' lives.

But diversity in terms of rivalry is not only relevant for the comparative evaluation of research. As Lakatos says, "*the sooner competition starts, the better for progress*" (Lakatos 1976a, 69). The common saying goes "competition is good for business" and also in science, being confronted with the successes of a rival theory, method or perhaps medical product, an established line of inquiry may be put "on the spot" (Brown 2001, 185) and the people involved may be spurned on to increase or renew their efforts. The epistemic point is indirectly connected to this argument: to be able to improve our current claims and theories, we may first need to evaluate them and find out what is still lacking. This idea of diversity in terms of beneficial competition raises another important issue for the criterion of productiveness: So far, I have talked about productiveness in terms of the contribution of the SoR to societies' needs at a particular point in time. The idea of rivalry or competition is about more than such a snapshot, however. We can also think of productivity in terms of the long-term expectations we have: Even if at the moment the available research is not as helpful as it could be, perhaps the current distribution will lead to many important discoveries in the future. This kind of reasoning is also particularly wide-spread with defenders of basic research, which, while maybe not directly contributing to societies' needs, is supposed to provide many long-term benefits.

There are thus multiple instrumental arguments for preferring a diverse SoR. Still, I will focus on productiveness in terms of efficiently preventing relevant epistemic and practical risks, instead of directly referring to diversity here. Diversification, after all, remains a heuristic that may not always be beneficial. For example, as Kitcher himself admits, if "one strategy, S_1 is much more promising than all the others [...] it will be best to put all the eggs in one basket" (Kitcher 2011, 194). Sometimes, calls for more diversity of approaches may in itself be a waste of resources and lead to alethic risk. Authors such as Oreskes and Conway (2012) or Biddle and Leuschner (2015) have argued that science critics may use claims of imbalance—and lack of diversity—as a way to cast unfounded doubt on scientific findings or at least forestall their acceptance in society. With the alleged pause in global warming, we have encountered an example for this idea that is connected to imbalance in the SoR. The general, intuitive worry that if the SoR is not diverse, it might be suboptimal, is exactly the kind of unexplained feeling that I want to account for by explicating the worries about cases of imbalance in the SoR. In the end, for each case of imbalance one will have to argue that the distribution at hand makes things worse in terms of alethic or practical risk in order to argue that there actually is a problem.

5.2.2 The Criterion of Productiveness

From the previous discussion, a framework that is supposed to account for cases of imbalance in the SoR on grounds of productiveness needs to include the following subordinate criteria (cf. Table 5.1):

1. Alethic Risk	
1.1 Deviation from True Belief:	**Error**: The risk of recipients of the SoR forming false beliefs. **or** **Uncertainty**: The risk of recipients of the SoR being uncertain about true propositions.
1.2 Ignorance:	The risk of recipients of the SoR being unaware of questions or decision options.
2. Practical Risk	
Practical Error:	The risk of recipients of the SoR making practical decisions which fall short of the goals of their judgment.
3. Efficiency	
The productiveness of a distribution of research increases the less resources are needed to achieve the same level of risks reduction.	

Table 5.1: Explaining imbalance in terms of productiveness

I have argued that imbalances opposed to productiveness can be understood in terms of the risk that recipients of the SoR may be led to erroneous inferences.[10]

10. A remark about the concept of risk is in order: Risk is generally understood as some probability

I have distinguished two main types of risk: **alethic risk**—that is, the chance that inferences might be problematic in terms of truth or true belief—and **practical risk**. Alethic risk was further differentiated into the risk of agents inferring mistaken beliefs from the SoR—what may be called **alethic error**—, lesser forms of diverging from the truth—what Goldman calls ignorance and I propose to dub **uncertainty**—and, lastly, **ignorance**—defined as the risk of recipients being unable to take relevant questions or options into account when making their inferences because they are not aware of them. Practical risk, on the other hand, is about **practical error**, that is, errors in decision making. Rather than about deviating from truth, practical error is about outcomes which do not satisfy the ends that the decision was supposed to serve. In the context of toxicity assessment, for example, those ends may include the competing goals of safeguarding the public from exposure to toxic substances and, at the same time, not inflicting undue economic damages onto the producers of those substances.

But even if some distribution of research is equally risky as another, one of them may be perceived as a problem with productiveness while the other is not. This is because productiveness is also about the **efficiency** of science: if a distribution of research offers very little return in terms of risk reduction[11] but requires lots of resources, we may consider it to be a case of research waste.

Note that all risk is relative to relevance. In this chapter, I discuss problematic imbalance in terms of the contribution of the SoR to a given question to be answered or a decision problem to be optimized according to an existing set of goals. Only risk which is **relevant** in the sense that it may affect inferences that could contribute to the

of harm. While sometimes, especially in retrospect, we may assume that some imbalance definitely was harmful, imbalance will often be about some uncertain negative consequence. This may be either because we do not know what inferences actually will be drawn, and with what background-knowledge—which is especially important in connection with miscommunication—but also because we often cannot know what the outcome of a neglected type of research would have been. This is especially clear in the case of undone science, where we may sometimes estimate, but can never know for certain if the alternative lines of inquiry would have changed anything about the inferences to be informed by the SoR.

11. It may appear counter-intuitive that I explicate productiveness in negative terms, that is in terms of risk, and not in terms of knowledge gained or right decisions made. However, the criterion of productiveness is supposed to be used to explain why we consider certain distributions of research to be a problem, that is, it is supposed to mainly apply to negative examples.

given question or decision-problem also affects productiveness. Different subsidiary questions may be relevant to the main question to different degrees; the more relevant an answer to a question is, the greater the risk of being wrong or ignorant about the answers.

I have made yet another distinction in this chapter that is of general interest; however, it is not about *why* it is a problem but *how* it operates, and thus, how we may alleviate its effects: On the one hand, I described cases where—like in the concept of misleading claims—risk is caused by **miscommunication**. In those cases, the audience of the SoR is likely to take some feature of the distribution of research to imply something that is not justified: for example, the public might interpret the amount of research on the alleged pause in anthropogenic global warming to imply that this warming indeed has stalled, or doctors and patients might take the fact that there are no published negative results for some medical intervention to imply that there are no doubts to be had about its effectiveness, while this impression would change if they were informed about the existence of publication bias. If communication concerning the SoR were improved, or so the idea, the public might—rightly—be less skeptical about global warming or—rightly—more skeptical about certain drugs or treatments.

On the other hand, I have argued, there are cases that constitute "more substantial issues" (p. 125). But what exactly is meant by that? Obviously, they are more substantial in the sense that the risk caused by these issues cannot be avoided by better communication alone: I have claimed that, for example in the case of undone science, we might hope to make the public aware of some lacuna in the SoR, but we cannot communicate the results of some research that has not been done. That is, undone science is also about a **lack of knowledge** about the outcomes of research.[12] When we lack knowledge about the existence of some possible but very relevant line of inquiry also ignorance of questions and options of interest may lead to problems with inferences further down the line.

12. In a case of pure practical risk—cf. subsection 5.1.2—, however, it would not be a lack of knowledge, but the mere fact that the research doesn't exist which causes the problem.

5.2.3 The Alethic, the Practical and the Ethical

In the preceding section, I have distinguished alethic and practical risks as possible undesirable outcomes of impediments to productiveness. I have also already hinted at a further distinction within practical risk: we can distinguish ethical risk from merely instrumental risks. This difference maps onto the distinction between Kitcher's notion of significance and what I have called relevance:

Relevance, in my framework, runs parallel to what Kitcher, in his description of significance graphs (cf. section 2.3), called "significance in terms of" some ulterior goal. Kitcher's concept of "significance" is more normatively laden than my notion of relevance, however. Relevance is always relative to some given interest, but it does not depend on this interest being morally justified. Significance, on the other hand, indicates something that is not only desired but also actually desirable:

> Think of a problem for investigation as arising when some entity of a specified type is sought. Problems worth pursuing can be labeled as significant. Those problems are adequately solved when an item is produced that is close enough to the type sought to serve the purposes that confer significance to the problem. (Kitcher 2011, 105)

Significance thus only accrues to scientific projects and claims when they contribute to problems actually worth pursuing, not just to any problem somebody might have. The formulation of productiveness in terms of knowledge needed by the citizens is open towards the question if this need is descriptive—knowledge which supports whatever goals the citizens might have—or an ethical criterion along the lines of Kitcher's significance. If science is unproductive in terms of significance, it constitutes a moral problem; if is unproductive in terms of relevance to any goal, it is just instrumentally problematic, and the gravity of the problem hinges on the importance of the goal. I prefer the second concept—the concept of relevance—because it is more open, and allows us to explain intuitions about imbalance even when they only refer to some interest or goal we do not share are consider worth pursuing. In the case of ignored genes, for example, it is just assumed that "our understanding of human biology and disease, and provide new targets for drug discovery" (Edwards et al. 2011, 163) should be the goal of inquiry without providing a moral rationale; in the case of regulatory paradigms, the authors themselves seem to avoid making any normative judgment of the effects of undone science, but mention "threats to wildlife and humans from persistent, toxic, industrial chlorinated pollutants" (Frickel et al. 2010, 449) as the issues which excited the "extensive citizen activism" (ibid., 449).

With the concept of relevance alone, we cannot say that these truly are problems all things considered: while there appear to be threats human health or the environment which are caused by imbalance, one could also say that the careers of scientists or the profits for the chlorine industry do perhaps thrive on the SoR. And while in these contrast we might have an intuitively clear answer for what is ethically preferable, other cases may not be as clear-cut: in energy-research, for example, is what we should aim for providing renewable energy for society such as wind- or solar-power, or do we focus on providing carbon-neutral power as fast as possible, which would include nuclear power, but also comes with side-effects that we might find undesirable? Judging if a SoR is truly problematic will, in the end, sometimes rest on answering these very difficult questions.

But how would we even begin to provide an answer? In Kitcher (2001), the significance of problems seems to rest on some member or fraction of the public assigning importance to it, with the discussions of the ideal deliberators in the process of well-ordered science serving as a filter for morally undesirable interests and for balancing the interests existing in society. However, it is notoriously difficult to say anything substantial about the outcomes of such an ideal and thus hypothetical process beyond what has been invested in the set-up: In Kitcher's case, this means that what should count as a significant process is supposed to be sanctioned—or rather, sanctionable—by deliberative, representative, democratic procedures. In other works such as Kitcher (2011) or Kitcher (2015), he ties scientific to ethical progress, and conceives of the latter in terms of "remedying altruism failures" (Kitcher 2011, 47); an idea which in turn is connected to Kitcher's views on the original function of ethics. Kitcher claims that these are not supposed to be about the actual history of humanity. They just outline one possible way in which ethical practices might have arisen (cf. ibid., 43-45). Perhaps this is to avoid the impression that his focus on "the evidence available to us (evidence from psychology, primatology, archaeology, anthropology, and evolutionary theory)" (ibid., 43) commits him to a form of naturalistic fallacy; however, it is then even less clear how his account is supposed to ground an ethical theory. I will, therefore, not discuss this idea in any more detail. Generally, I do not recommend any particular ethical theory which would allow judgments about the significance of research goals or decision-problems. There is a wide variety of ethical theories which have philosophical, political, and societal support, and I do not want to restrict the recognition of problems with imbalance to only one particular standpoint. However, I still want to point out some general features of moral arguments that can support criticisms of the SoR that are based on the criterion of productiveness I have presented in the preceding section.

Firstly, the kinds of arguments that can substantiate problems with productiveness as moral problems need to be consequentialist: Productiveness is supposed to be judged on the effects produced by the research, in terms of risks to either—in the case of alethic risk—epistemic values, or—in cases of practical risk—negative impacts on decision making. But this does not imply any particular form of consequentialism, that is, I do not commit myself to hedonism or any other form of utilitarianism. For example, one could also talk about threats to productiveness in terms of putting human rights at risk, which arguably is the case when—as in many of the cases discussed—the research-goals under discussion are about human health. Secondly, the criterion of productiveness emphasizes *efficiency*. That is, it focuses on producing a maximum amount of useful research output while using a minimum of resources. An ethical theory behind productiveness must thus not only be able to account for the minimization of risk, but also also for the proportion of risks avoided to research resources expended. Thirdly, the discussion so far has been impersonal, that is, it has focused on the effects on society as a whole, and disregarded the problem of how the benefits provided by science should be distributed among different societal groups. While the latter is an important question, especially if we recognize that the needs of individual members of society may be very different, and often in conflict, it is not the focus of the criterion presented in this chapter. In the next chapter, therefore, I will pay respect to ethical positions that include a demand for fairness and discuss a criterion of epistemic justice, which is clearly distinct from the concerns about productiveness.

Chapter 6

Imbalance and Injustice

In the previous chapter, I have discussed the criterion of productiveness as a way to explain what is problematic about certain cases of imbalance in the SoR: the SoR is in imbalance if—by way of alethic or epistemic risk—the focus on or neglect of a particular type of research makes it less likely that science will efficiently provide answers to significant questions. I mainly focused on the question of what makes some distribution of research problematic given that we have already determined what questions should be pursued; that is, I have focused on productiveness in terms of relevance. I have also explained that, if imbalance is to be judged ethically problematic, we need to consider productiveness in terms of significance. This comes down to maximizing the good produced by scientific research, which implies a connection between productiveness and consequentialist ethics. But is the criterion of productiveness and the maximization of goods the only connection between imbalance and ethical considerations? Are there other criteria central to ethical theories that we need to explain why we find certain cases of imbalance problematic? In this chapter, I will argue that being able to explain the examples of imbalance in the SoR also requires an account of justice. In the first section, I focus on the example of neglected diseases and argue that, while both productiveness and justice may often be connected in cases of imbalance, they are often not clearly disambiguated, and the aspect of fairness has undeservedly received less attention by the authors involved in the discussion. In the second section, I turn towards the debate about epistemic injustice in social epistemology, and distinguish three concepts of injustice which may apply to imbalance in the SoR. In the third and last section, I summarize these findings in a framework for explaining imbalance in terms of injustice.

6.1 Neglected Diseases & Justice

In meta-ethics and general ethical theories, detailed reflections on the role of science are sparse. And in applied ethics—for example, in research ethics—the perspective on the SoR is mostly ignored. There are, of course, quite a few contributions about individual problematic cases—for example, the ones included in chapter 3—and some more systematic accounts of how research is supposed to be organized, such as Kitcher's "well-ordered science" (Kitcher 2001, 2011). However, there is very few literature comparing several ethical theories to problems with the SoR. As one of the few exceptions, De Winter and Kosolosky (2014) analyze cases which constitute problems with the distribution of research and make use of five different ethical theories to argue that these cases "are problematic on ethical grounds, showing that they are moral failures" (ibid., 701). The authors focus on a set of three different problems with the research agenda, two of which have also appeared in chapter 3 of this dissertation: firstly, the example of neglected diseases discussed above; secondly, the problem of a lack of unpatentable research; and thirdly an example about mainstream agricultural research versus agroecology. After presenting these cases, they discuss several prominent ethical positions and how they might apply to the examples. They conclude:

> Whether one is a utilitarian, an adherent of Rawls's theory of justice, a human rights advocate, an adherent of Kitcher's ethical theory, or a classical liberalist, the conclusion seems to be the same: the distorted research agendas in the health sciences and the agricultural sciences are morally problematic. (ibid., 723)

While this analysis is a valuable attempt to conceptualize these cases, in the following I will argue that the authors' interpretation in terms of maximization of goods overshadows an equally important alternative explanation in terms of justice or fairness, obscuring important differences between both kinds of argument. For the purposes of this dissertation, I need to provide a list of criteria that allows us to give a comprehensive and coherent account of the intuitions to the effect that some distribution of research is problematic. The range of ethical positions discussed by De Winter and Kosolosky do include all relevant aspects in principle; however, their account glosses over a crucial difference between the two major criteria, which is needed to explain these intuitions: productiveness as the maximization of epistemic goods, and justice as a matter of their distribution.

This difference already becomes apparent in the first example discussed by De Winter and Kosolosky (2014): the "problem of neglected diseases" (ibid., 703). Based on two earlier papers (Flory and Kitcher 2004Reiss and Kitcher 2009), the example can be summed up in the worry that research on some diseases which cause an extraordinary amount of suffering and primarily affect the poor—main examples include Chagas disease, malaria and others (ibid., 265)—are neglected in favor of less severe afflictions which primarily affect the affluent. One possible mechanism behind this phenomenon is connected to the fact that "[r]esearch dollars come almost entirely from the wealthy part of the world, and the suffering from malaria, tuberculosis, and a large number of infectious agents happens elsewhere" (Flory and Kitcher 2004, 40), and, "public R&D funds of high-income countries, which have the largest budgets at their disposal, are primarily allocated to research that is tailored to their own health interests" (De Winter and Kosolosky 2014, 704). Another explanation of how neglected diseases come about refers to the commercialization of research and the relationship between investments and expected profits:

> Drug development is very costly, and thus only chemicals for which there's a large potential market will be chosen for research and development. [...] For our purposes, neglected diseases will be those that multinational companies ignore on the grounds that, however many potential buyers there might be for a future drug, the overall revenue accruing would be too small to meet the constraints of profitability. (Reiss and Kitcher 2009, 265)

De Winter and Kosolosky first consider a very simple version of utilitarianism—which they link to Bentham's account—according to which we should evaluate actions based on maximizing pleasure and minimizing the pain produced. Starting with this basic account, they give a first explanation for why neglected diseases can be considered morally problematic: the amount of suffering avoided—and thus, happiness augmented—by producing research aimed at combating diseases which mainly affect people in developing countries would be much greater than the impact of research which is done on the health problems of people in wealthier societies, such as diabetes or high blood pressure; but in fact, the actual distribution of research is inverted (cf. De Winter and Kosolosky 2014, 706).

Compare this account with the second ethical theory De Winter and Kosolosky use to analyze the example of neglected diseases: John Rawls' Theory of Justice. Here, the authors' arguments proceed from the idea that the affected parties, under the conditions of Rawl's original position—ignorance of "particular facts", "their own place in society" and "their own conception of the good" (ibid., 712)—, would not endorse the cases under discussion as just or fair. Concerning neglected diseases, they conclude:

> So we can expect more people to be able to achieve their life goals if more resources would be allocated to health research for the poor, and less to the development of medicines for health conditions for which effective treatments are already available. So for parties in the original position, who do not know whether they are rich or poor, such a reallocation of resources would increase the probability that they can achieve their aims. Therefore, parties in the original position cannot rationally accept the current allocation of resources in the health sciences and the corresponding research agenda. This means that this agenda is, according to Rawls's theory of justice, unjust. (ibid., 713)

While I agree with the authors that neglected diseases can be criticized both from the perspective of classical utilitarianism and from the perspective of a Rawlsian conception of justice, there is an important difference in both how it would be criticized and what, exactly, would be the object of criticism: in short, the utilitarian critique is about maximizing happiness and objects to neglected diseases as an inefficient distribution of research; the Rawlsian critique, on the other hand, is about a procedure that is supposed to guarantee an impartial judgment of principles of justice, and objects to neglected diseases as an unfair distribution of research. Rawls' theory of justice proceeds in analogy to a maximin principle (cf. Rawls 1971, 150-161): According to Rawls—and in contrast to the version of utilitarianism sketched above—we should not only think about maximizing goods. The maximin principle for rational choice under conditions of uncertainty asks us to prioritize the best worst possible outcomes. Similarly, what—according to Rawls—can justify even an unequal distribution of goods is that, when we imagine ourselves as the least advantaged of society—the worst case for who we might turn out to be once the veil of ignorance is lifted—the redistribution still is beneficial. The explanation in the quote above, however, is not about putting ourselves in the shoes of the worst off; it is about alternatives to neglected diseases leading to "more people to be able to achieve their life goals" which for people in the original position would "increase the probability that they can achieve their aims". This, however, is still about an impersonal maximization in terms of the proportion of people being able to achieve their aims. A genuine argument in terms of justice, in contrast, would have to criticize that there is *inequality* concerning who is able to achieve their life goals, and that, if we consider the people who are worst off—the sufferers in less well-off parts of the world, already stricken by poverty—, they are even further marginalized by their afflictions being ignored by science.

While both lines of criticisms coincide in this particular case—that is, the phenomenon of neglected diseases may be considered both inefficient or unproductive and unfair—, in others, they could come apart. Even if they don't, it is worthwhile to point out that there is more than one aspect of the case that is problematic. Therefore, we need to clearly distinguish these lines of argument to provide a comprehensive framework for analyzing cases of imbalance.

But this is not a difficulty in De Winter and Kosolosky (2014) alone. The difference is already muddled by the discussion in the original papers which dealt with the concept of neglected diseases: There, the criterion which allows the authors to pick out a disease as neglected in the required sense is the "concept of a disease's fair share of research resources" (Flory and Kitcher 2004, 41) and the corresponding "'fair-share'-principle: at least insofar as disease problems are seen as comparably tractable, the proportions of global resources assigned to different diseases should agree with the ratios of human suffering associated with those diseases." (Reiss and Kitcher 2009,

263) What the authors lament is that, for some diseases like malaria, the money spent on trying to combat the disease does not adequately reflect the suffering caused—measured in “the number of years of life lost because of the disease” (Flory and Kitcher 2004, 44). However, while “fairness” in ordinary language is usually applied to the treatment of people, the fair share principle, as cited above, is about giving a disease a fair share of the research, i.e., it is about the proportionality of the suffering caused by the disease and the resources spent on trying to deal with it. In the end, this fair-share principle corresponds to a utilitarian or at least consequentialist account, which is in line with the concept of productiveness I presented in the last chapter. While a Rawlsian account may agree with a principle of productiveness in some cases, it is precisely the differences to such a criterion of maximizing some good—in this case, useful research—which make it an account of justice. As outlined above, Rawls' conception does include the maximization of goods, but only insofar it is to the benefit of the least advantaged members of society. Consider his second principle of justice:

> Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) attached to offices and positions open to all under conditions of fair equality and opportunity. (Rawls 1971, 83)

Now imagine a case in which there are two possible new medical research agendas under consideration: the first focuses on the development of new drugs that are very costly to produce but extremely effective in combating a type of disease which causes great suffering in all parts of the world; the second focuses on alternative interventions which are much less costly—perhaps research into diets or exercise—but where it seems they might also be much less effective; perhaps alternative approaches only slightly alleviate the consequences of the type of disease in question, while the drugs that would be developed under the first alternative are likely to fully cure patients that receive them. Assume further that the first agenda will produce treatments which are only available to people in areas of the world with strong healthcare systems, because the cost of production and the lack of infrastructure in other places does not allow the new drugs to be distributed there. The alternative lines of research, however, would be applicable everywhere. For the sake of argument, let us imagine that the overall suffering reduced—once again perhaps in life-years gained—is greater for the first agenda than for the second. Now, the difference between the utilitarian argument and the argument from justice should be clear. If we favor productiveness alone, the first agenda will appear preferable; if, however, we subscribe to the difference principle—the a)-part of Rawls' second principle of justice above—we have to label the first agenda unjust, as it prefers a treatment which maximizes the effectiveness of interventions over the benefit to the least advantaged members of society—the sufferers in less-well of parts of the world who would not have access to the costly new drugs. While this example is a toy case used to demonstrate the important difference between the two lines of ethical argument, it is modeled to fit one of the other examples of imbalance in the SoR appearing in chapter 3 but also in De Winter and Kosolosky (2014): unpatentable research. In the previous discussion of the example, I have assumed that we might criticize a lack of unpatentable research because the focus of the latter "could well be", as Brown puts it "a far superior treatment, both cheaper and more beneficial" (Brown 2008, 197). Similarly, also De Winter & Kosolosky claim that "sometimes, non-medicinal solutions are more effective than medicines" (De Winter and Kosolosky 2014, 714). However, especially when we discuss science policy and real-world cases, whether unpatentable research is preferable depends on very difficult judgments about the likely effectiveness of different interventions and, perhaps more importantly in this context, who stands to benefit most from the future research. The toy example also resembles what S.D. John (2014) discusses as the "'prevention paradox': 'population strategies' that reduce the (relatively) low risk of many can be more effective at improving overall population health than 'high risk strategies' that reduce the (relatively) high risk of smaller subpopulations" (ibid., 28). While not strictly about justice in the sense of this chapter, his analysis shows that similar worries about how to capture our intuitions about interventions connected to uncertain risks to different groups of society are very relevant for discussions about population health. Trying to account for all cases of imbalance in the SoR in terms of a benefit in overall productiveness or by the fair-share principle alone obscures some of the difficult choices to be made about which probabilities to assume, and sometimes, which ethical values we give priority: Do we focus on the overall contribution to our significant problems produced by our research, or do we first want to make sure that nobody has an unfair advantage or disadvantage from some SoR?

But even in examples like the problem of neglected diseases, where it is obvious that the current situation is untenable according to both criteria, we would do well to distinguish the two lines of justifying moral criticism. Firstly, an ethical framework that is supposed to expose the problematic nature of a phenomenon should closely capture the intuitions connected to the cases we want to account for. Furthermore, while both a critique from justice and from productiveness may explain *that* we might intuitively believe that neglected diseases are a moral problem, only focusing on productiveness obscures parts of the *why*. Only by including a justice-based account can we capture the intuition that the case is not just about an ineffective use of our research resources, but also about the wrong of favoring a particular group of people over another, especially if this is to the disadvantage of people who are vulnerable and discriminated against even without this imbalance in the SoR. Also, even if one of these arguments may be enough to claim that a case is morally problematic, the force of this criticism may be bolstered by giving two relatively independent arguments to the same effect.

Still, one may object that the need to distinguish productiveness and justice may disappear if we use the right metric for determining what is productive, or more precisely, what problems are significant. There is some plausibility in arguing that part of what makes neglected diseases so problematic in terms of the fair-share principle is that people in the less well-off countries are more vulnerable to disease. Therefore, the same affliction may cause more suffering for this group of people than for others; or, on the flip side, by dealing with disease in poorer countries, we might be able to do more good than elsewhere:

> Effective technology for eliminating malaria in Africa might thus serve as a basis for ameliorating other forms of suffering. Plainly, if such socio-economic considerations were incorporated into a refined conception of a disease's fair share, they would only increase the gap that divides fair share from actual expenditure. (Flory and Kitcher 2004, 47)

When confronted with the criticism that maximizing the utility of the many come at the cost of diminishing the utility of few excluded people—that is, with the accusation that utilitarianism may endorse injustice—utilitarians can refer to the principle of diminishing marginal utility: providing some amount of good to those who had few or nothing before may produce more utility than providing the same good to one who is already well-off. Another ten euros a month—to reiterate an old example—, for someone who has almost nothing and might thereby have the opportunity to afford another few meals may be a huge improvement, while for a millionaire another ten euros—certain thresholds aside—may bring no additional utility at all. However, there are also cases where this explanation does not work, and Rawls theory of justice may be seen as a reaction to these deficits of utilitarian theory:

> Yet this[—the principle of marginal utility—]is only a contingent matter. If some people are very adept at turning resources into well-being – they are so-called "utility monsters" – then a utilitarian should support a rule that privileges them. This seems repugnant to justice. As Rawls famously put the general point, "each member of society is thought to have an inviolability founded on justice which....even the welfare of every one else cannot override" (Rawls 1971, p. 28; Rawls 1999, pp. 24–25). (Miller 2017)

This can be applied to my starting point of neglected diseases: it may seem plausible that the suffering we may be able to alleviate by focusing on diseases affecting people in developing countries is greater than what we may be able to achieve when focusing similarly grave afflictions of people in the industrialized world. But how much people really suffer does not solely depend on their situation but also on internal factors like the subjective perception of the disease and the tolerance for one's circumstances. While the authors in the debate try to provide some more objective measures—such as the years of life lost as a consequence of suffering from the disease, or "the discounted value of years of life lived with disability" (Flory and Kitcher 2004, 44)—these measures do not reflect diminishing marginal utility. Establishing any universally agreeable standard that allows us to argue that some people necessarily suffer more from a comparably terrible disease than others because of their circumstances is a hard task. But even if we were able to show that, by defining significance in the right way, a theory focusing on the productiveness of science alone may be able to accommodate intuitions about injustice, the effort needed would show that injustice is a topic that in itself deserves our attention—we would just have managed to include the criterion as a sub-aspect of productiveness.

In the remainder of this chapter—just like with productiveness in the last—I will, therefore, provide a closer look at justice as a criterion for evaluating the imbalance of SoRs. There are different subordinate aspects of justice, some of which may be applied to some cases of imbalance, but not to others.

6.2 Imbalance as Epistemic Injustice

With Rawls' theory of justice, we already have one candidate for a conceptual resource that may allow us to characterize certain cases of imbalance as problems with injustice. But while there has been an ongoing ethical discussion of Rawl's account for decades now, much more recently an active discussion about justice has also begun in social epistemology. In the following, I approach the criterion from the perspective of this debate about epistemic injustice, mainly because of the relevance of this concept to research as an epistemic good. Also, in this debate, there is still much instability in what is to be considered part of the concept of epistemic injustice, and what should be excluded; by applying it to cases of imbalance and discuss the differences between some of the examples, I will contribute my own take on the taxonomy.

6.2.1 Distributive Epistemic Injustice

The current discussion about epistemic injustice can be traced back to Miranda Fricker's *Epistemic Injustice: Power & the Ethics of Knowing*. In this book, the author presents a very detailed and insightful account of two novel forms of injustice: The first one is testimonial injustice, where the "basic idea is that a speaker suffers a testimonial injustice just if prejudice on the hearer's part causes him to give the speaker less credibility than he would otherwise have given" (Fricker 2007, 4). For an obvious, drastic example, consider "the case where the police don't believe someone because he is black" (ibid., 4). The second form of injustice Fricker calls hermeneutical injustice: "the injustice of having some significant area of one's social experience obscured from collective understanding owing to a structural identity prejudice in the collective hermeneutical resource" (ibid., 155). Here, a central case "is found in the example of a woman who suffers sexual harassment prior to the time when we had this critical concept, so that she cannot properly comprehend her own experience, let alone render it communicatively intelligible to others." (ibid., 6) I will relate both cases to the examples for imbalance in the SoR below. Before that, however, some notes on the general relationship of epistemic injustice and imbalance are in order. Fricker's concept of injustice is especially interesting because it is supposed to be about inherently epistemic injustices; both central cases are about someone being "wronged specifically in her capacity as a knower." (ibid., 20)

However, there are also some possible points of contention between Fricker's account and the explananda in this dissertation. Think back to the central example of an intuitively unfair SoR above: the phenomenon of neglected diseases. This case is about who stands to benefit from medical research and the claim that people in industrialized countries stand to gain more from the knowledge produced than people who suffer from disease in developing countries. It is very clearly a case which is about distributing a basic epistemic good, that is, the knowledge produced by medical research. Fricker, however, has been very vocal about the claim that her concept of epistemic injustice is not supposed to be about distributive justice. This is closely linked to her understanding of what makes her cases intrinsically epistemic problems:

> Given how we normally think about justice in philosophy, the idea of epistemic injustice might first and foremost prompt thoughts about distributive unfairness in respect of epistemic goods such as information or education. In such cases, we picture social agents who have an interest in various goods, some of them epistemic, and question whether everyone is getting their fair share. When epistemic injustice takes this form, there is nothing very distinctively epistemic about it, for it seems largely incidental that the good in question can be characterized as an epistemic good. By contrast, the project of this book is to home in on two forms of epistemic injustice that are distinctively epistemic in kind, theorizing them as consisting, most fundamentally, in a wrong done to someone specifically in their capacity as a knower. (ibid., 1)

In the following, I will argue that Fricker's initial attempt to outright exclude issues of distribution from an interesting concept of epistemic injustice is not convincing.

I will, however, also point towards one important aspect of Fricker's cases, which cannot possibly be described as a wrong in terms of distribution. There has been an ongoing debate about this point between David Coady (2010, 2017) and Fricker (2007, 2013, 2017). Agreeing with Alvin Goldman (2003), Coady convincingly argues that also distributive epistemic injustice is distinctively epistemic:

> I think that Goldman is right that interesting true belief is an intrinsic value, which is neither reducible to any other value nor plausibly seen in entirely instrumental terms. Hence, Fricker is wrong to think that questions about the just distribution of this epistemic good are only incidentally epistemic. (Coady 2010, 112).

There is no good reason to believe that the matter of fairly distributing knowledge is not intrinsically an epistemic issue. It may be objected that my account does not only concern the distribution of the epistemic goods themselves—that is, the products of research and the knowledge which can be gained from relying on them—but also fairness concerning the benefits in terms of applications of this knowledge: The example of neglected diseases, for example, is not about people in less well-off parts of the world having less access to knowledge than other groups of people. It is about the effects of diseases from which they suffer, which could be at least partly remedied by an increase in research effort. This, someone could argue, is not an intrinsically epistemic problem, but rather a problem with unfair distributions of health resources. But how would we measure inequality in terms of knowledge alone? As I have argued at multiple occasions throughout this dissertation, we should not concern ourselves with just any truth, but only with relevant and significant truths. Consequently, also a distribution of knowledge or research products should reflect not only the quantity of knowledge available, but how important these epistemic resources are for the problems of the people relying on it. This, then, means that epistemic injustice is always connected to the value of those goods in terms of the public's interests. Note that, as I have argued in subsection 4.2.1, also purely epistemic interests—think curiosity—are admissible here, and can thus contribute to the value of research.

Still, it may be criticized that justice in the distribution of knowledge is thus only about the distribution of an instrumentally, not intrinsically valuable good. However, this is also true for other goods such as economic wealth. As Coady argues, there is a parallel between epistemic justice and justice in the distribution of wealth or political power; it would be very strange, however, to claim that the distribution of wealth is only incidentally an economic issue (cf. ibid., 105), or that the distribution of political power is only incidental to political theory. The remaining difference between Coady's and Fricker's account of epistemic justice, on this view, would mostly be about the former being about unjust ignorance or error, while the latter would be about an unjust deficit in credibility or intelligibility. Coady has defended this view against Fricker's arguments to the effect that testimonial injustice cannot be a distributive injustice because credibility is no finite good (Fricker 2007, 17-21; Coady 2017, 63-64). I agree with Coady that we can, in principle, understand testimonial injustice as injustice in the distribution of credibility. This, however, is something very different than injustice in the distribution of knowledge. As I will argue below, these differences do set Fricker's concept of testimonial injustice apart from cases which are about distributive epistemic injustice in terms of knowledge: they should be understood as a form of participatory epistemic injustice.

But what exactly distinguishes Fricker's cases from justice in the distribution of knowledge? Fricker has since acknowledged Coady's criticism and admitted that "the unfair distribution of epistemic goods such as education or information is an important kind of social injustice in its own right, and may often be closely intertwined with the discriminatory kind" (Fricker 2013, 1318). As the quote suggests, however, she still wants to distinguish Coady's concept from her own, now emphasizing that the latter is about discriminatory epistemic injustice, which means that the wrong is supposed to be tied to prejudices against or the marginalization of certain societal groups (cf. Fricker 2017, 53). But prejudices and marginalization can just as well be part of injustice in the distribution of knowledge: consider the problem of gender imbalance in the philosophical syllabus (cf. Saul 2013, 44-45), which may be rooted in—perhaps unreflected—prejudices about women authors being less important to the history or SoR in philosophy. At the same time, it would seem to me that much of what is specific to testimonial and hermeneutical injustice would also still be unfair if it were not caused by prejudices or marginalization: if, for example, the statement of the victim of a crime would be ignored by the police or the court because the police officer or judge disliked them personally and wanted to degrade them, this would perhaps be less of a wrong than if it were caused by systematic discrimination; however, much of Fricker's analysis of the case would still apply.

What then, if anything, does differentiate Fricker's cases and distributive injustices concerning knowledge and education? As far as hermeneutical injustice goes, I am skeptical if there is truly a non-distributive aspect: in the end, the case of the unintelligibility of harassment experiences—one of the prime examples for hermeneutical injustice—and the case of neglected diseases are both about a group of people being unfairly disadvantaged by a lacuna in a collective cognitive resource, only that the first is about a lack of hermeneutical resources, while the latter is about a lack of medical knowledge. Both are problematic primarily because this injustice keeps them from a good which would enable them to deal with problems significant to them. What then makes hermeneutical injustice special is mainly in what this inability amounts to, namely the powerlessness to communicate one's experiences or even to process and understand them for one's own sake.

A very clear distinction can be made, however, between distributive forms of epistemic injustice and a specific interpretation of testimonial injustice: the latter is not only a wrong in *instrumental* terms, that is, in terms of the unfair effects of being excluded from an epistemic process on the results of said process. Rather, what is special about Fricker's testimonial injustice is that it also constitutes an *intrinsic* epistemic wrong: the wrong of being denied one's deserved epistemic status, and thus being degraded as a knower. I will further explore this non-distributive aspect—which is essential to Fricker's work—in the second half of subsection 6.2.2. Before that, however, I will go into more detail about distributive epistemic injustice: What, exactly, does it mean that epistemic goods like knowledge are supposed to be fairly distributed?

In the following, I will rely on the account of Faik Kurtulmus and Gürol Irzik, who have provided a very insightful analysis of justice in the distribution of knowledge. They define the problem in the following way:

> Our central claim is that justice requires that people have the opportunity to

> acquire knowledge about matters that they have an objective interest in as individuals and citizens, and this in turn requires that the epistemic basic structure of their societies produce and disseminate such knowledge and provide them with the capabilities they need for assimilating it. Accordingly, a systematic lack of opportunity to acquire knowledge one needs as an individual and a citizen to reason about the common good, her individual good and pursuit thereof because of the way the epistemic basic structure of her society is organized is an injustice. (Kurtulmus and Irzik 2017, 129-130)

There are at least three important aspects in their account of epistemic justice that should be highlighted:

Imbalance and structural injustice. Firstly, their analysis is about structural injustice. Injustice is not traced down to individual actions of persons in isolation, but to the institutions which shape the process of knowledge production. Fricker has acknowledged that also her version of epistemic injustice is not just about individual actors, but that also the rules of institutions can, for example, deny certain people the right to give testimony (cf. also Fricker 2007, 56-57). Already in Mill's criticism of the sorry state of religious freedom in his times, he describes a structural testimonial injustice against atheists:

> This refusal of redress took place in virtue of the legal doctrine, that no person can be allowed to give evidence in a court of justice, who does not profess belief in a God (any god is sufficient) and in a future state; which is equivalent to declaring such persons to be outlaws, excluded from the protection of the tribunals; who may not only be robbed or assaulted with impunity, if no one but themselves, or persons of similar opinions, be present, but any one else may be robbed or assaulted with impunity, if the proof of the fact depends on their evidence. (Mill 2015, 31)

As I have said in chapter 3, the concept of imbalance in the SoR is primarily supposed to be about structural problems. The SoR itself is always only the final product of a long chain—or rather, an intricate web—of complicated interactions between many different people, whose transactions are constrained and reinforced by the rules and makeup of scientific institutions. This starts, of course, already on the level of individual studies: the outcomes, publication, and reception of which depends on the structure of the community. The possible effects of an individual personal interaction on balance at the level of the SoR are even more limited. Anderson's account of structural epistemic injustice very precisely captures this:

> These lessons apply to epistemic justice as much as to distributive justice. Answering a complex question, or interpreting some significant phenomenon, typically requires that we elicit epistemic contributions from numerous individuals and connect them appropriately. The cumulative effects of how our epistemic system elicits, evaluates, and connects countless individual communicative acts can be unjust, even if no injustice has been committed in any particular epistemic transaction. Nor can we count on the practice of individual epistemic justice to correct for all of these global effects. Rather, the larger systems by which we organize the training of inquirers and the circulation, uptake, and incorporation of individuals' epistemic contributions to the construction of knowledge may need to be reformed to ensure that justice is done to each knower, and to groups of inquirers. (Anderson 2012, 164-165)

Returning to Kurtulmus and Irzik (2017), the authors point towards three areas connected to the distribution of knowledge which are of relevance to justice: The production of significant knowledge, the dissemination of the findings to the public, and the opportunity for the public to make use of the findings, mostly in terms of a fair access to the education needed (cf. ibid.). Intuitively, imbalance in the SoR may be most strongly connected to the production of findings—often philosophers of science mainly discuss issues of justice in connection with the research agenda, as we have seen with Kitcher's work. However, imbalance in the SoR can also be about an issue with dissemination: excluded or inaccessible research is at least primarily about cases where research of the neglected kind does exist, but it is not published or accepted as a relevant contribution to the SoR. In chapter 5, we have also seen the importance of imbalance as a communication problem, where alethic or practical risk arises from misunderstandings about the implications of the published findings. I argued that overemphasis on the alleged pause in anthropogenic global warming, for example, could be less worrying if publications had been more careful about the terminology concerning the “pause”. Such issues of how results are communicated can also be seen as problems with the dissemination of knowledge. The issue of access to adequate education for making use of scientific results seems to be less a problem for the SoR; at least in the examples of imbalance I provided in this dissertation there seem to be no cases that can be interpreted as a problem with an unfair lack of education which arises because of a lacuna or overabundance of a type of research. Therefore, I will also focus on the distribution of knowledge as the main basic good, and not—like Coady does—also on education.

The structural interpretation fits the cases of imbalance which we may intuitively consider problems with injustice: in the case of neglected diseases, it is the interaction of the funding structure in medical research—the origin of most medical funding coming from affluent nations and the financial interests of the pharmaceutical industry—which leads to the diseases in question being neglected. Two examples which we will revisit in the next subsection are structural in the sense that institutional guidelines might lead to injustice: in the case of air-monitoring standards, the policy-framework that determines what environmental data is admissible shapes the imbalance in the SoR; with the case of standards of EBM, an inner-scientific evidence-hierarchy may be thought of as unfair to certain groups of people.

Imbalance and primary versus private goods. Returning to Kurtulmus’ & Irzik’s concept of injustice in the distribution of knowledge, the second aspect I want to highlight is their requirement that the knowledge to be fairly distributed among the members of society is supposed to be “about matters that they have an objective interest in as individuals and citizens” (ibid., 129). While the explicit connection to the citizens’ or individuals’ interests can easily be accommodated by the notions of relevance and significance introduced before, it is more difficult to incorporate the concept of *objective* interests into this framework.

Already Kitcher’s concept of significance does contain an element of interactive objectivity (cf. Douglas 2004, 463-465): the initial preferences of members of society are to be filtered through the tutoring provided by scientific experts. This is to avoid the consequence that, otherwise, the unfiltered preferences “would favor short-term practical inquiries over research of long-term significance, that the emergent research agenda would be myopic and probably unfruitful" (Kitcher 2011, 112). Also, the procedural requirements of representativeness and mutual engagement are supposed to improve and aggregate the initial interests and thereby reflect Kitcher's idea of an underlying ethical project, which comes down to "remedying altruism failures, concentrating on those occasions on which members of the group were thwarted in obtaining those things that 'made their lives go well' [...]" (ibid., 50). Also, Kitcher excludes certain ethically problematic projects by positing that the ideal deliberators should set up constraints of inquiry, "always stemming from the recognition that a particular way of pursuing inquiry would violate the rights of some individual or group" (Kitcher 2001, 121).

Kitcher's requirements on legitimate interests are procedural; Kurtulmus and Irzik (2017) offer a more substantive account of what interests should be admissible in questions about epistemic injustice. Following the Rawlsian idea that distributive justice should concern primary goods only (cf. ibid., 141), they argue that the justice in the distribution of knowledge should be about those interests which are connected to "questions that bear on individuals' plans of life and the common good" (ibid., 132), help to "ensure a well-functioning democracy" (ibid., 131) and, more specifically, contribute to "just legislation" (ibid., 131). Excluded from their account of distributive epistemic justice are private interests, such as "a factory owner who wants to build a gadget at minimum cost [and thus] has an interest in finding technologies that will decrease production costs" (ibid., 134). I will use a less restrictive concept because I would argue that also these other cases, where private profits from public knowledge are unevenly distributed, can constitute injustices—just considerably less grievous ones. However, I keep the distinction between private goods and more, general, basic interests to account for some of the intuitions concerning examples of imbalance in the SoR: In cases like neglected diseases or unpatentable research, for example, the private interests of people involved with the pharmaceutical industry—mainly, the interest in increased profit—may lead to a deficit in basic goods such as the patient's health. When, like in those cases, a private and a basic good are pitted against each other, the latter should weigh much more than the former. With Kitcher, we could even say that, in these cases, structural features of medical research violate a basic human right to health. The differences between the goods or interests at stake should also be considered when arguing that, because some people have contributed more to the scientific process, they should also receive more benefits from the SoR. It is, for example, not implausible to argue that pharmaceutical companies have to be able to make some profit from investing huge amounts of money into drug development, and should thus also profit more from the research than others. However, with the distinction just introduced in mind, we should consider if we can set up the "*epistemic basic structure*" (ibid., 129) of society in a way that this justified demand for profit does not violate other people's rights. There is an active discussion about the right way to deal with commercial interests in science (cf. Carrier 2008; De Winter and Kosolosky 2014; Brown 2017; Reiss 2017) that contains proposals for institutional reform such as to "establish a medical prize fund to reward medical innovation (partly) on the basis of the impact of this innovation on global health" (De Winter and Kosolosky 2014, 6). Also, some mechanisms to support research which may balance industrial interests are already being tested. For example, in Italy, a fixed percentage of the money used by pharmaceutical companies for promotional purposes has been redirected towards independent research on, among other issues, rare and thus neglected diseases (cf. AIFA 2010, 75).

Imbalance and access to knowledge. The third and last aspect of the account concerning justice in the distribution of knowledge by Kurtulmus and Irzik (2017) I want to point out is that, actually, their discussion is not about knowledge per se, but about access to knowledge:

> This qualification is needed for two reasons. First, there is no guaranteed way of acquiring true beliefs. Our most reliable way of acquiring true beliefs is through well-conducted research, and yet it is not a guarantee for acquiring knowledge; it is merely our best bet. Second, what justice requires is not that people know all the facts that they have an interest in, but that they can come to find out about them, or rather the results of well-conducted research on them. (ibid., 130)

However, these qualifications do not seem particularly strong to me; the authors use the shorthand "distribution of knowledge" without any inconsistency throughout most of the text. I mention their notion of "opportunity to acquire knowledge" mainly to point out that this concept of justice in terms of fair distribution of opportunities is not the same as equality of opportunity as referred to in Rawls' principles of justice that I quoted before (cf. Rawls 1971, 83); Kurtulmus' and Irzik's take on injustice is not about being denied equal access to particular positions or offices—or in, our context, the opportunity to fulfill certain epistemic roles or functions. In the following subsection, I will introduce participatory epistemic injustice as a criterion that reflects this other aspect of equality of opportunity.

6.2.2 Participatory Epistemic Injustice

Let us now return to Fricker's concept of testimonial injustice. While it can be understood as a form of distributive injustice, it is not primarily about injustice in the access to knowledge as a resource, but about a credibility deficit. In the following, I will clarify the epistemic function of what is distributed here; that is, what capacities as a knower are diminished by this kind of injustice. I will argue that it can be considered a special form of a broader concept of procedural injustice, which I will call—following Heidi Grasswick (2017)—participatory epistemic injustice. I will then explain that this form of injustice can constitute two different kinds of ethical wrongs: an instrumental wrong, where it is about being denied the opportunity to influence the outcomes of the scientific process, and an intrinsic one, where it is about the exclusion of social groups as a form of epistemic degradation.

Fricker herself already distinguishes a primary and a secondary aspect of harm: a first, which is intrinsic and consists in the lack of credibility assigned to the speaker itself, and a second, which refers to the extrinsic consequences of this denial (cf. Fricker 2007, 44-48). The primary aspect can also be described in terms of the second principle of justice which we encountered earlier in this chapter: it comes down to the fact that "the subject is wronged in her capacity as a **giver** [my emphasis] of knowledge" (ibid., 44). This we can understand as the sufferer being denied equal opportunity to contribute to an epistemic practice; it can be seen as a violation of condition (b) of the second principle of justice: "Social and economic inequalities are to be arranged so that they are […] (b) attached to offices and positions open to all under conditions of fair equality and opportunity." (Rawls 1971, 83) Here, "equality" should not be taken to mean that everyone should always be given the opportunity to provide to epistemic practices to the same degree, i.e., among other things, that the credibility assigned to each person should always be the same. Instead, the level of credibility assigned should match the trustworthiness of the person in question, or, in Fricker's words: "Epistemological nuance aside, the hearer's obligation is obvious: she must match the level of credibility she attributes to her interlocutor to the evidence that he is offering the truth." (Fricker 2007, 19) What is meant by equal opportunity, then, is that this judgment of trustworthiness or evidence be made without prejudice, and everybody who may have something to contribute, within reason, is given due consideration. Exclusion in this sense is an important aspect of several cases of imbalance in the SoR. However, I will be using it in a broader sense than just testimonial injustice. As I have argued concerning public trust in science, the range of people whose actions determine what ends up as part of the SoR goes far beyond contributions in terms of testimony: what to fund, what to pursue, what and how to argue, how to present one's findings, what to criticize in reviews, what to publish how and what to deem relevant for inclusion in the SoR are all questions which determines the distribution of research. People may be included or excluded in all of these decisions. Being wrongly denied the opportunity to contribute to the scientific process as a patient or a citizen with local expertise—perhaps closest to Fricker's concept of testimonial injustice—, being denied the opportunity to try and pursue a scientific career, or being denied the opportunity to shape the goals of research through participating in the funding and approval process can all be seen as being wronged in the capacity to contribute to the epistemic process that is scientific research, to be denied a place in the collective societal search for answers to significant questions. Exactly this point is made by Grasswick (2017):

> Testimonial injustices are crucial to understanding the unjust impediments to the central epistemic activities related to knowledge transmission, yet epistemic injustices can also afflict many other core epistemic activities concerning the generation of knowledge itself. [...] When this happens as a result of systematic forces of oppression, a participatory epistemic injustice results. (ibid., 315-316)

It may be objected that this concept of access to and participation in epistemic activities goes beyond Rawls' demand in terms of offices and positions—being included in science may not come with some official position, after all. However, Rawls' use of these terms mainly reflects his general concept of justice as principles of justice for institutions:

> The primary subject of the principles of social justice is the basic structure of society, the arrangements of major social institutions into one scheme of cooperation. [...] The principles of justice for institutions must not be confused with the principles which apply to individuals and their actions in particular circumstances. [...] Now by an institution I shall understand a public system of rules which defines offices and positions with their rights and duties, powers and immunities, and the like. (Rawls 1971, 54-55)

The concept of participatory injustice presented here goes beyond official positions; however, it is consistent with Rawls' account in that it is about a major social institution—the institution of science—and how this system assigns duties and powers throughout the scientific process.

Several of the examples for imbalance in the SoR can be understood as cases of participatory epistemic injustice: A structural testimonial injustice we can see in the case of air-monitoring standards, where citizens and activists are denied the opportunity to contribute their bucket data to the studies of air quality. Similarly—although this requires some additional assumptions not necessarily part of the original description—if the standards of EBM exclude research which is based on the individual experiences of patients, this also denies this group of people the possibility to act in their capacity as givers of knowledge. And indeed, epistemic injustice involving patients has received explicit attention in the recent "epistemic injustice"-literature (cf. Kidd and Carel 2017; Carel and Kidd 2017; Fricker 2017, 85-89). We can, therefore, define a second category of epistemic injustice, which is conceptually distinct from injustice in the distribution of knowledge. Following David Miller (2017)), we may interpret this distinction along the dimension of procedural versus substantive justice, that its, as a distinction "between the justice of the procedures that might be used to determine how benefits and burdens of various kinds are allocated to people, and the justice of the final allocation itself." (ibid.) Participatory injustice, then, is a form of procedural injustice where some group or individual, because of implicit or explicit prejudices, are denied the opportunity to contribute to epistemic processes.

Saying that participatory injustice is about being denied the opportunity to contribute does not, however, clearly point out what is ethically wrong about these cases. We can distinguish two underlying criteria. The first is participatory injustice as an instrumental wrong. This idea corresponds to Fricker's secondary aspect concerning the harm of testimonial injustice:

> Turning now to the secondary aspect of harm, we see that it is composed of a range of possible follow-on disadvantages, extrinsic to the primary injustice in that they are caused by it rather than being a proper part of it. They seem to fall into two broad categories distinguishing a practical and an epistemic dimension of harm. (Fricker 2007, 46)

Equality of opportunity here can be seen as a means to an end, where the end is a fair distribution of the basic goods affected by this opportunity: in our case, primarily the research—in terms of knowledge or other cognitive resources. Not being able to participate bars those affected from influencing what will or will not end up in the SoR, increasing the alethic and practical risks for those individuals or groups when their perspectives and inputs are disregarded. In the court example, not being able to testify may lead to me being convicted of a crime I didn't commit; in the example of air-monitoring standards, the inability of citizens and activists to contribute to the SoR and thus the connected policy decisions may negatively impact the health of people living in the vicinity of the industrial plants in question. Interpreted as positive criteria, participatory and testimonial justice are procedural values, which are about ensuring or at least make it more likely that substantive injustice in the SoR can be avoided.

This instrumental reading of participatory and testimonial injustice is, however, not Fricker's main concern. She highlights the intrinsic wrong of being exposed to this form of injustice:

> When someone suffers a testimonial injustice, they are degraded qua knower, and they are symbolically degraded qua human. In all cases of testimonial

> injustice, what the person suffers from is not simply the epistemic wrong in itself, but also the meaning of being treated like that. (Fricker 2007, 44)

Fricker argues that testimonial injustice literally adds insult to injury: not only is the person being denied the opportunity to contribute robbed of the possibility to influence the outcome of the inquiry, but being denied also comes with the insult of being marked as someone who should not be allowed to contribute. It assigns them a lower status as a knower than those whose input is included.[1] Note that the aspect of dehumanization emphasized by Fricker may not be as prominent in some cases that I interpret as participatory epistemic injustice. In some cases the impression of degradation is strong: gender imbalance may be the result of implicit bias, connected to the prejudice that women are not apt to become scientists or philosophers, denying them the rationality or "genius" which is sometimes deemed a prerequisite (cf. Leslie et al. 2015); when evidence standards lead to the exclusion of the experiences of patients, they perhaps may feel that they are marked as irrational and thus dehumanized (cf. Kidd and Carel 2017). However, in cases like the air-monitoring example, what is denied to the citizen scientists or activists, in contrast, is not a basic human capacity, but the recognition of their expertise. While we may also see an insult in that—we can speculate that locals may feel insulted by the idea that they may have less valuable insights into their local environment than the scientists do—there is also no general insult to someone "qua human" in the assumption that the training and knowledge scientists possess may sometimes privilege their contributions over those of laypeople. An element of degradation may thus be attached to participatory injustice in various degrees. With this form of injustice there is also finally an aspect of Fricker's account which can hardly be understood as an issue of distribution: being insulted or degraded is not about having less access to a good, where, because there is not enough respect

1. It seems that for Fricker, there are actually three kinds of harm in testimonial injustice: the intrinsic epistemic wrong, the connected wrong of degradation, and the secondary, instrumental wrong. It is not quite clear what besides degradation Fricker sees as the intrinsic epistemic wrong; Ronald Dworkin (2002) suggests "three kinds of participatory consequence: symbolic, agency, and communal" (ibid., 187). The "symbolic" wrong can be likened to the wrong of degradation, "agency" to what Rawls (1971) calls "realization of self" (ibid., 84), and "community" is connected to the value of being part of and strengthening a collective. While these further aspects are interesting, their relation to imbalance is similar to that of degradation, so I will not pursue them here for reasons of simplicity.

to go around, some people are unfairly insulted more than others. Injustice, here, has to be used in a broad sense, not in the sense of fairness but in the sense of a violation of someone's right do respect and dignity, which is unjust if it is not redressed. Still, it is a distinctly epistemic form of injustice, as it is about being denied a particular *epistemic* status.

Still, I am not convinced that we should, like Fricker seems to suggest, pick out this aspect to generally demarcate cases like the ones used to introduce testimonial injustice from cases of distributive justice: Testimonial injustice is connected to both the wrong of epistemic degradation and the extrinsic effects on distributive epistemic justice. In many cases of imbalance in the SoR, the latter even is more prominent in the discussion. Also, there is another connection between intrinsic participatory injustice and the distribution of knowledge we have not yet considered. Grasswick (2017) introduces the idea of epistemic trust injustice:

> Laypersons need to trust scientific communities in order to benefit from the very best and most relevant scientific results along with the scientists' professional judgments of the status of scientific research, including its uncertainties. With this necessary epistemic role of trust comes the possibility of what I call epistemic trust injustices. Epistemic trust injustices occur when, due to the forces of oppression, the conditions required to ground one's trust in experts cannot be met for members of particular subordinated groups. (ibid., 319)

Grasswick's ideas imply that cases of participatory injustice can also indirectly lead to distributive injustice. Indirectly, that is, because what disfavors the excluded groups is not a lack of influence on the distribution of research. They suffer a distributive injustice because, due to the participatory wrong, they have reason to believe that science lacks competence, benevolence, or integrity, and thus to reject the scientists' results—which, in turn, implies that they cannot benefit from the research at all. Whyte and Crease (2010) discuss a similar idea as a part of what they call "unrecognized contributor cases" (ibid., 415). They recount the example of sheep farmers' local insights being excluded from the research activities of scientists from the British Ministry of Agriculture, Fisheries, and Food (MAFF) investigating the effects of the Chernobyl disaster on agriculture (cf. ibid., 415-417) and conclude:

> The unrecognized contributors then tend to acquire distrust of the credentialed experts. These actors are unable to see the value the experts' models and methods given that the latter were unwilling to engage local experience and knowledge. Though, as previously mentioned, such lay contributions may not improve all aspects of the scientific analysis, the fact that they are outright rejected by scientists breeds distrust, regardless of whether we can precisely foresee the improvements. This distrust is not conducive to resolving scientific controversies like that between the MAFF scientists and sheep farmers. (Whyte and Crease 2010, 417)

It should be noted that both the example in Whyte and Crease (2010) and epistemic trust injustice are about problems with warranted trust rather than trustworthiness. Warranted trust can be diminished in at least two ways: firstly, by the trusted being actually not trustworthy, and secondly, by the trusting party not having a warrant for this trustworthiness. This allows for cases in which science at one point did historically have a problem with integrity or benevolence as far a particular group is concerned, but conditions have since improved. I such cases, science may already be trustworthy as it is; still, scientific institutions need to engage with the groups in question to make their change in disposition visible to the public, and thus provide a warrant for trust. The idea of epistemic trust injustices is, of course, no blanket justification for distrusting science, even though most people will probably be able to point out some cases where scientific institutions did not work in an entirely benevolent way. The amount of distrust should reflect the gravity of the insult, and should be limited to the areas of science where it originated; otherwise, we would have to admit that science skeptics—outright refusing, for example, to vaccinate their children—would probably be justified in their positions. The lesson to be learned, however, is that such cases may not always be only about how reliable the scientific evidence is, but that we also have to consider the history of science in its relation to various social groups and personal biographies if we want to remedy these problems.

I do, however, not believe that we should establish epistemic trust injustice as an independent category or wrong of injustice. What would it mean for a criterion to be independent in the required sense? Substantive injustice can still exist even if the processes determining the SoR is procedurally just; even if all relevant social groups are represented in the scientific process, we might still end up with a SoR that serves the goals of particular groups better than others. As a short excursus, take Kitcher's concept of well-ordered science: It is an inherently procedural idea, asking us to approximate a process of ideal deliberation. Very roughly speaking, well-ordered science is about ideal deliberators who, in a democratic process of mutual engagement and tutored by scientific experts, decide what research projects to support with how many resources, how to conduct the research, and, in the end, how to apply the results (cf. Kitcher 2001, 118). What supposedly makes this process fair to the social groups affected by the research is a requirement of representativeness on the views held by the ideal deliberators:

> [T]he ideal procedure attempts to incorporate the views of every member of the pertinent society. It's an open question as to whether the collection of ideal deliberators contains distinct idealized representatives of each citizen or whether we can assume that people divide into groups whose members are sufficiently similar that they can be represented en bloc. (ibid., 123)

This is supposed to avoid problems of substantive distributive injustice:

> A group is inadequately represented when the research agenda and/or the application of research results systematically neglects the interests of the members of that group in favor of other members of society. (ibid., 129)

However, Kitcher admits that even the highly idealized procedure of well-ordered science does not guarantee substantive distributive justice in all cases. Consider the worst possible outcome for the stage where the deliberators decide on a list of significant problems:

> Finally, if the intersection of the sets of lists deliberators accept as fair turns out to empty, collective preferences are determined by vote on all candidates drawn from the union of these sets of lists. (ibid., 119)

In this scenario, even if all the requirements on procedural justice have been met, some of the groups and people represented by the ideal deliberators may be outvoted in this final step, having to accept a list of significant projects that they deem unfair.[2] We can also think of cases where the requirements of procedural justice are not met, but still, the outcome is substantively just. As Dworkin (2002) puts it in the context of political equality:

> For a benevolent tyranny, in which none of our assumptions about democracy held, might nevertheless produce a just property scheme, and might otherwise respect the distributional goals of the right conception of equality; indeed it might produce a more egalitarian distribution than a democracy could. But no tyranny could advance the participatory goals any egalitarian community would also aim to secure. (ibid., 187)

Similarly, also a benevolent science administration which respects everyone's interests might perhaps achieve a distributively just distribution of research, but still, be unjust in terms of intrinsic participatory injustice.

Such independence, however, is not given in the case of Grasswick's epistemic trust injustice. What exactly is the wrong underlying that concept? At face value, it seems to be about a substantive, unfair deficit in access to epistemic resources, where certain groups do not trust science and thus cannot profit from scientific expertise. But this is not just about cases where there is no relevant knowledge which could in principle be beneficial to the wronged group or where the distribution of research lends itself to epistemic or practical risks for that particular group. While in that case, this would also constitute a reason not to trust science, it would also already be a case of substantive distributive injustice in terms of knowledge and need no additional explanation in terms of trust. No, what Grasswick wants to point us towards are cases where the research itself is reliable, but science can still not be trusted because of some injustice which has happened in the past:

> Trustworthiness is situated. In the case of a subjugated group that has experienced a history of oppression, a preponderance of evidence against the epistemic

2. In cases of imbalance in the SoR, there is another reason to fear that well-ordered science may fail to provide a just distribution: if the imbalance pre-dates the deliberation process, one of the cornerstones of well-ordered science—the tutoring by scientific experts (cf. Kitcher 2001, 119-120) and also the "*atlas of scientific significance*" (Kitcher 2011, 127)—may itself be skewed in problematic ways.

> trustworthiness of scientific communities (leading to responsibly-placed distrust rather than responsible trust) can result when those scientific communities have participated in and contributed to that very history of oppression. In such circumstances, an epistemic trust injustice occurs, wherein members of the group are unable to satisfy the conditions of responsible trust. (Grasswick 2017, 319)

The concept of epistemic trust injustice is inevitably tied to such historical instances where a disposition of scientists or scientific institutions that violated the requirements of integrity or benevolence towards the social group became visible. But such instances are already in themselves instances of injustice—perhaps in terms of degradation or insult—and it is those events which make epistemic trust injustices unfair. Otherwise, it would also be quite difficult to say what the injustice is: An unfair distribution of trust? An unfair distribution of trustworthiness? Is the current situation in itself degrading or insulting to the social group? Neither of these options seems to capture what Grasswick describes. The wrong of epistemic trust injustices is just a consequence of the original wrongs, and should thus not be considered an injustice in itself.[3].

6.3 The Criterion of Epistemic Injustice

I can now summarize the findings of this chapter and provide an expansion of the framework for explaining additional aspects of imbalance in the SoR (cf. Table 6.1): The major ethical problems in cases of imbalance beyond productiveness are matters of **epistemic injustice**. Epistemic injustice refers to the unfair treatment of groups and individuals in their capacity as a knower.

We can distinguish two major forms of epistemic injustice: first, substantive **injustice in the distribution of knowledge**, where members of certain groups have less access to knowledge as a resource—be it for satisfying their curiosity or their

3. While, in our context, it makes most sense to discuss this issue with warranted trust in connection with participatory injustice, past violations of scientific benevolence and trust could take many other forms: unethical research practices involving members of the affected social group as subjects, for example; or racist, sexist or otherwise discriminatory remarks made by persons with authority in scientific institutions. These other forms of degrading practices, however, cannot clearly be connected to issues with imbalance in the SoR.

practical projects. Throughout my discussion, I have referred to Rawls' account of justice. I do, however, not want to fully commit to his or any other very specific theory concerning justice, because the intuitions about examples of imbalance may not be rooted in any one such account. In spite of this, I have excluded a strictly egalitarian interpretation of epistemic justice. There are good arguments which can justify uneven distributions of knowledge: I mentioned a criterion of desert, where those who contribute more to science than others may also expect to receive greater benefits. Desert-based arguments can, of course, also be criticized for various reasons. For one, we have encountered the Rawlsian Difference Principle, according to which any unequal distribution must be "to the greatest benefit of the least advantaged" (Rawls 1971, 83). Also, such inequalities become especially worrisome if they trade off **knowledge as a primary good**—where it contributes to basic human needs such as health or other requirements codified in human rights, or ensures the functioning of democracy or our judicial system—against **knowledge as a private good**—where it only affects some additional personal gain. Some interests in knowledge may even be intrinsically opposed to the rights of other groups, and should thus be excluded from consideration. But there is also merit to the thought that without any advantage for private sponsors, private research might disappear altogether; and that without these private contributions, the situation also for those least advantaged by the SoR would be even worse. While these arguments rest on empirical assumptions and the details of particular cases and institutional arrangements, these considerations do speak against a wholly egalitarian conception of distributive epistemic justice. Therefore, where there is an uneven distribution of epistemic goods we may initially suspect an imbalance of the SoR in terms of distributive injustice; to substantiate this claim, however, we will need to consider further arguments about the specifics of the case.

Second, epistemic injustice may appear as **participatory epistemic injustice**. This concept describes cases where certain social groups are excluded from contributing to the scientific process, and this exclusion is not justified by admissible epistemic reasons but rooted in explicit or implicit prejudice. Participatory injustice is always procedural[4], but it can both appear as an instrumental and as an intrinsic wrong: on the one hand, deficits in equality of opportunity to contribute to the scientific process denies the wronged social group the possibility to influence its outcomes. Therefore, it can lead to substantive injustice in the distribution of knowledge: **instrumental participatory injustice as a deficit in epistemic influence**. Because this form of participatory injustice is wrong in terms of subsequent effects on the distribution of knowledge, it can be considered one form of procedural distributive injustice. On the other hand, being denied the opportunity to contribute to epistemic activities, where this exclusion is based on prejudices about the excluded people's epistemic capacities, also constitutes an intrinsic epistemic wrong; being excluded from the scientific process can be perceived as an insult, where the group in question is denied the required expertise or, in more extreme cases, the status of a rational human being: **intrinsic participatory injustice as epistemic degradation**[5].

Not only the influence deficit, but also the aspect of degradation itself, however, may lead to secondary injustices further down the line: if past interactions between scientific institutions and a social group have eroded the latter's belief in the competence, benevolence or integrity of science, they may lack **warranted trust**, and can thus not make use of scientific knowledge to the same degree as others. This additional instrumental wrong may both be about a continual lack of trustworthiness on the side of science, or about a communication problem where the social group has no warrant to trust scientific institutions but where they actually are trustworthy.

Participatory epistemic injustice, with its reference to **equality of opportunity**, leans more towards egalitarian conceptions of justice than distributive epistemic

4. Cases of imbalance in the SoR are primarily presented as structural problems. Therefore, I follow Kurtulmus and Irzik (2017) in also treating injustice in the distribution of knowledge as an institutional problem, i.e., of the epistemic basic structure of society. One could argue that, thus, also injustice in the distribution of knowledge should be considered to be a form of procedural injustice; it is about the process and structures which govern the distribution of research, after all. However, the wrong of injustice in the distribution of knowledge is always, in the end, evaluated in terms of the substantive differences in access to relevant knowledge. It can be reduced to the assessment of the outcome; therefore, considering it a procedural criterion as well would be misleading.

5. In subsection 7.1.2, I will argue that this last criterion is better understood as a problem with the process leading to the distribution, and not with the research distribution itself. Therefore, it does not explain imbalance in the SoR. However, I include the criterion here so I can point out this difference, and also because it is an important aspect of epistemic injustice in general.

injustice does; this, however, does not imply an **equality of influence**. Because of talent or acquired skills and knowledge, particular people might be more suited for certain functions within epistemic institutions. For example, ceteris paribus, a trained scientist may be able to provide better scientific policy advice than an interested citizen without this training. What participatory justice does demand of science, however, is to aim at equality concerning the opportunity to acquire the relevant attributes and at a fair evaluation of all those who want to contribute to the scientific process, no matter what social group they belong to.

Lastly, it should be noted that epistemic injustice can appear both as a structural injustice—where injustice is rooted in the explicit or implicit norms and interactions of various institutions and individuals—and transactional injustice—where the injustice can be traced down to the actions of particular individuals. While imbalance in the SoR could be about both kinds of injustice, the latter is less probable, and—as of yet—there are no examples of imbalance which constitute transactional epistemic injustices.

With this and the previous chapter, I have completed the analysis of the criteria for explaining imbalance suggested by the conception of public epistemic trustworthiness of science introduced in chapter 4. In the last part of the dissertation, I will combine these criteria and apply the resulting framework to the examples from chapter 3. In the remainder of the text, I will discuss the implications of these results for a concept of imbalance in the SoR and related notions such as independence and objectivity.

1.: Substantive distributive injustice	
The SoR is unjust if one or more social groups undeservedly have less access to knowledge that can support them in solving significant problems than other groups.	
2.: Participatory epistemic injustice	
The SoR is unjust if one or more social groups are denied the opportunity to contribute to the scientific process because of implicit or explicit prejudice.	
Two wrongs of participatory epistemic injustice	
2.1.: Instrumental participatory injustice (deficit in epistemic influence)	Being denied the opportunity to contribute can be instrumentally wrong because a lack of influence on the scientific process may lead to injustice in the distribution of knowledge. Constitutes a form of **procedural distributive epistemic injustice**.
2.2.: Intrinsic participatory injustice (epistemic degradation)	Being denied the opportunity to contribute can be wrong because it constitutes a form of degradation where it denies groups or individuals their deserved epistemic status.

Table 6.1: Explaining imbalance in terms of injustice

Part III

Results

Chapter 7

Explaining Imbalance in the SoR

7.1 Application of the Framework

The main goal of the second part of this dissertation was to provide a conceptual framework that can account for all the examples of imbalance in the SoR I presented in section 3.1 (cf. Table 3.1, p. 86). "Account", in this context, should be taken to mean that for each of the examples, the framework offers a criterion that explains why the case at hand is intuitively perceived as problematic. But how should we understand the idea that the criteria explain these intuitions, and how would we assess if they actually do? In the introduction, I have suggested that the overall approach used throughout this project can be reconstructed as a process of mutually adjusting the intuitive judgments about the cases and a systematic conceptual framework of imbalance that is rooted in philosophical background theories. In the course of the project, additional examples for similar problems connected to the SoR were added to the list, and, consequently, I expanded upon my initial ideas about what constitutes the problematic imbalance in those cases. This resulted in the reflections about the normative background, which form the second part of this dissertation. In each of the chapters, I concluded with an overview of a criterion that can be used to criticize a SoR as imbalanced: with trustworthiness as my starting point, I developed both an account of epistemic productiveness and an account of epistemic injustice, which together are supposed to provide an explanation for why imbalance is a problem in each specific case.

In this chapter, I show that the framework can indeed successfully explain imbalance in the SoR. First, I will argue that the framework is comprehensive. Comprehensiveness, here, is supposed to mean that for each of the examples presented in chapter 3, the framework provides an explanation for the intuitive judgment that there is a problem with the SoR—or more specifically, with lacunae or overabundance in the distribution of research.[1] If there would be cases that cannot be explained using the framework, then either some of the examples would not constitute imbalance, or the framework would be incomplete. As the concept of imbalance was introduced by means of the examples, the first option would have to be very well justified, while the latter would show a definite flaw in my analysis up to this point. But when can a criterion be said to explain an intuitive judgment about these cases? A corresponding explanation needs to answer to the question: "*Why* is the example at hand problematic?" That is, the criteria need to provide us with reasons for judging that the distributions of research described in the examples are undesirable. In subsection 7.1.3, for each of the criteria I will provide a general template for a type of argument that contains the criterion as a premise and concludes—directly or indirectly—that some distribution of research constitutes a problematic imbalance in the SoR. In subsection 7.1.3, I will provide an actual argument based on at least one of these templates for each of the examples. If an argument based on one of the criteria is convincing in a particular case, the framework has provided a reason for thinking that this case is a problematic imbalance in the SoR, and it has thus explained why we are justified in having an intuition to that effect.

Secondly, I will also claim that the framework is coherent. More specifically, I claim that I have provided an analysis of deeply interconnected problems that become visible from a particular perspective on science—the perspective on the SoR. It is not just a set of unrelated criteria that is based on a random list of examples constituting several independent issues with scientific research. Conceiving of imbalance as a coherent, rather than a disjunct set of problems, also allows me to use it to argue for a

1. This is not to say that all intuitions must always be explainable: it would theoretically be possible that none of the criteria can explain why a case seems problematic, and this is not because the framework, but rather the intuitions themselves are to be criticized. However, as seen below, this particular outcome never occurred.

specific concept of objectivity to which it is opposed. I establish the comprehensiveness of the framework in three respects: Firstly, in subsection 7.1.1, I will examine relations between the criteria that are a direct consequence of my theoretical discussion in the second part of the dissertation, largely independent from the application to the examples. Secondly, in subsection 7.1.2, I will operationalize the criteria, and show how the criteria and sub-criteria are inferentially connected to the conclusion that a specific example constitutes a case of problematic imbalance in the SoR. Thirdly, if there were disjunctive sets of examples, each explained by a single criterion alone, I would have been hard-pressed to argue that I have provided one framework explaining a shared phenomenon. But as I will discuss in subsection 7.2.1, the application shows that most examples can be explained by not one, but several overlapping criteria.

7.1.1 Relations between the Criteria

In chapters four through six, I have introduced three normative concepts that we can apply to SoRs and the scientific institutions that produce them: epistemic trustworthiness, epistemic productiveness and epistemic injustice. Trustworthiness is not itself a criterion for explaining imbalance, i.e., a lack of trustworthiness is not a reason for judging a distribution of research to be problematic; rather, imbalance in the SoR is a reason for not trusting science. In chapter 4, I have argued that, when we ask what generally makes science worthy of societies' trust, epistemic productiveness, and epistemic justice are two main requirements. In chapter 5 and chapter 6, I have argued that these two concepts are indeed criteria for imbalance in the SoR: Not all problems with epistemic productiveness or justice are problems with imbalance. For example, if the evidence collection in some area of science were methodologically unsound, this would surely be a problem for productiveness, but would not necessarily lead to lacunae or overabundance of some type of research. Still, violations of these two criteria *are* at the heart of the examples of imbalance given in section 3.1. In the following, I will show that there are inferential relations between all three concepts, demonstrating the framework's coherence on a theoretical level.

Two relations are already suggested in the above: both epistemic productiveness and epistemic justice were introduced as requirements for epistemic trustworthiness. Therefore, if science is unproductive or unjust, science is also not trustworthy.

What about productiveness? Consider, once again, Grasswick's (2017) idea of epistemic trust injustices: the idea that, because of some past injustice on the side of science, members of certain societal groups will not be able to trust scientific institutions, and thus also be unable to profit from the products of research. This idea shows two possible further relations: Firstly, if the public is unable to trust science, the research will also be unproductive for members of society. Note that "being able to trust" is not the same as science being trustworthy. In reality, who trusts whom might rest on a variety of historical and psychological factors—people might trust information which is not trustworthy, and distrust trustworthy research. This also holds where only ideal rational actors are concerned: While trustworthiness should be a requirement for making use of scientific research, we still have to demand that those actors know of the trustworthiness of science—there should be warranted trust. Secondly, the concept of epistemic trust injustice also hints at an indirect connection between productiveness and justice: if science is not just, then it is also not trustworthy at least for some members of society, and thus, less productive.

Furthermore, the same issues that make a distribution unjust may also make it less productive: If certain people or groups are excluded from the scientific process—and where this is not done so because of an unprejudiced judgment of their epistemic capabilities—this constitutes a participatory injustice. But also, where specific views and inputs are excluded, science may become less diverse in the sense of a lack of epistemically fruitful competition between methods and points of view. As James Robert Brown (2001) puts it:

> The way to ensure the optimal diversity of rival theories is to make sure we have a wide variety of theorists. Currently, the pool of those who make the conjectures is heavily skewed toward white males from upper-middle-class backgrounds living in wealthy countries. Future hiring must change the proportions so that a larger number of females, minorities, and others with different biases are included in the group of theorists. In short, affirmative action is needed for the sake of improving the growth of knowledge; pluralism for the sake of epistemology. (ibid., 187)

I do not claim that increasing the diversity of scientific personnel is the only or even the best way to ensure an environment that can also lead to an increase in epistemic productiveness; also, who and which groups and categories should be recognized as a minority that needs more representation is not uncontroversial. What I want to argue here is just that it is plausible that ensuring that diverse people have a say in the scientific process *can* improve the output; thus, in some cases of participatory epistemic injustice, productiveness *may* suffer as a consequence. Note that this is not the same claim as the one that—in line with the idea of instrumental participatory injustice—access to knowledge for a particular group can suffer, and thus lead to distributive epistemic injustice. Rather, this reflects the argument that, because of a lack of rivalry and competition, overall productiveness may decrease.

I now have examined the relations between productiveness and trustworthiness as requirements for trustworthiness, and between trustworthiness and justice as contributing to productiveness. But how do trustworthiness and productiveness affect justice?

From my previous discussion of the three concepts, there is nothing to indicate that an overall increase in either trustworthiness or productiveness contributes to distributions of research being just. We can, however, rephrase the criterion of substantive distributive injustice in a way that refers to productiveness: a SoR constitutes a problem with substantive distributive injustice when the current distribution of research is less productive for members of some particular group than for others and we can give no reason for why this inequality is justified. That is, if a SoR is fair, then this does imply something about the distribution of productiveness among members of society. However, it does not imply that productiveness has to be high; it may also be low, as long as the distribution puts no one at an unfair disadvantage. Because this is so, also trustworthiness interacts with justice: where, as in the idea of epistemic trust injustices, trust is unequally warranted for different societal groups, this may also lead to a distribution of productiveness that is unfair. Beyond that, however, my theoretical discussion suggests no further contribution of trustworthiness to the fairness of a SoR.

I have now discussed all six possible directions of connections between the criteria, but what kind of relations are these? Some I have called requirements, so one might be tempted to think about them in terms of necessary or sufficient conditions; then, the relationship between two such criteria would be that of logical implication[2]. But this interpretation of the term “requirement” is not intended here. I have said that trustworthiness is required for productiveness, and justice is required for trustworthiness. If we were to interpret this in terms of implications, this would entail—as per transitivity—that a SoR can only be productive if it is just. Which, on the flip side, would imply that there could be no SoRs that are productive but unjust. But we can think of counter-examples: it may be particularly productive to focus the research agenda on topics which are of interest to the scientists—perhaps because they then are especially motivated for their work. It may be most productive to work on projects that are of interest to the wealthy members of society, because then they might be more likely to provide funding. This might mean however, that some issues which are of low priority for those groups are neglected, and some people have to suffer from distributive injustice.

Trustworthiness, productiveness, and justice are not binary criteria; they come in degrees. A SoR can be more or less trustworthy, productive, or fair. Also the relations between them should thus be not understood as simple logical operators. This start already with epistemic productiveness and epistemic justice as requirements for the public epistemic trustworthiness of science, where the latter is defined in terms of the former. Those relations only indicate that, theoretically, the maximum amount of trustworthiness could be reached only if the SoR were maximally productive and maximally fair. But in actuality, because of resource constraints, there are only so many possible ways to distribute our resources among different types of research, and we might have to trade off a certain level of productiveness against some level of justice to arrive at the actual maximum in terms trustworthiness.[3] I have also already used possibilistic or probabilistic terms to describe some of the relations, e.g., when I say that participatory epistemic injustice may increase productiveness, but it also

2. This is because a necessary condition can be represented as the consequent, and a sufficient condition as the antecedent in an implication.

3. This is not supposed to indicate that the end-goal should always be trustworthiness; some might say that our principal goal should be productiveness; some, and I think we should interpret a Rawlsian maximin principle that way, would argue that justice should come first.

may not. The exact modal nature of the relations between the criteria may differ, and in the absence of empirical investigations, it is hard to give more precise analyses of the degree to which they support each other. However, the above discussion shows that already from my conceptual discussion, there are different kinds of mutual support-relations between the criteria, less strong, perhaps, between trustworthiness and justice. This, then, supports my claim that imbalance in the SoR, rather than an accidental grouping of unconnected issues, is a coherent concept referring to a set of interrelated problems.

7.1.2 Argument Templates

In the previous section, I have examined the theoretical implications for the relations between the three main concepts, which constitute the framework for explaining cases of imbalance. In this section, I will provide a list of templates that interprets the criteria and sub-criteria in terms of arguments. Operationalizing them in this way forces me to commit to a precise version of each criterion, so their applicability to the individual examples can be evaluated in subsection 7.1.3. Note that all these arguments are supposed to provide pro tanto or ceteris paribus reasons for some distribution of research constituting an imbalance, not all-thing-considered judgments. What appears as an undesirable distribution of research from one perspective—for example, because it leads to some practical risk—may turn out to be the best possible distribution overall—perhaps because all the less risky alternatives constitute particularly grave epistemic injustices. All things considered judgments will have to be based on the relative weight of the applicable arguments for imbalance where such conflicts arise.[4] As this possibility is explicitly foreseen, such conflicts would not indicate that the framework itself is inconsistent.

If the criteria are supposed to provide explanations for our intuitions that a particular example constitutes a problem with the SoR, in the end, they will have to provide support for the following main claim:

> **C:** The distribution of research—*a*—constitutes a problematic imbalance in the SoR.

4. However, as it will turn out in the next subsection, no obvious conflicts between the criteria became relevant in the examples of imbalance.

In the previous chapters of the dissertation, I have provided two principal lines of argument for such a claim. The first includes all arguments from productiveness:

	A1: Arguments from productiveness
P 1	If distribution of research x is less productive than possible alternative y, then x constitutes a problematic imbalance in the SoR.
P 2	a is a distribution of research & b is a possible alternative to a.
P 3	a is less productive than b.
∴	Distribution of research a constitutes a problematic imbalance in the SoR.

A second basic template describes arguments from epistemic injustice:

	A2: Arguments from epistemic injustice
P 1	If distribution of research x unfair, then x constitutes a problematic imbalance in the SoR.
P 2	a is a distribution of research.
P 3	a is unfair.
∴	Distribution of research a constitutes a problematic imbalance in the SoR.

In the chapters describing these two main criteria, I have also discussed more specific reasons for criticizing distributions of research. These do not constitute independent arguments for the same main claim, however; instead, they support these main arguments on multiple levels, forming an inferential web which can be represented as an argument map (see Figure 7.1). Both main argument templates have the same simple structure: premise one spells out the criterion as a general principle in the form of a universally quantified conditional. In premise two, the relevant individual constants for the particular example are introduced. Then, in premise 3, it is claimed that the "if"-condition of premise 1 is fulfilled, which allows us to conclude that the "then"-part is true as well.

For the application in the next section, it will be premise 3 that is the most questionable: can it be convincingly argued that the distribution discussed in the example is unfair, or that there is a possible alternative that would be more productive? In this, there is also the major syntactical difference between the two general argument templates: as I have argued in chapter 5, productiveness is always evaluated in relative terms; relative, that is, to some alternative distribution which would be more productive than the one at hand. This, of course, requires us to specify what this possible alternative could look like—see premise 2 in A1.

For arguments from productiveness, this means that they need to conclude in the statement that some distribution of research *a* is less productive than a possible alternative *b*. There are three major ways in arguing for this claim. The first is a type of argument based on the idea that *a* leads to more alethic risk than *b*:

	A1.1: Arguments from alethic risk
P 1	If distribution of research *x* leads to an increase in alethic risk over possible alternative *y*, then *x* is less productive than *y*.
P 2	*a* is a distribution of research & *b* is a possible alternative to *a*.
P 3	*a* leads to an increase in alethic risk over *b*.
∴	Distribution of research *a* is less productive than possible alternative *b*.

The second is based on practical risk instead:

	A1.2: Arguments from practical risk
P 1	If distribution of research *x* leads to an increase of errors or inaction in relevant decisions over possible alternative *y*, then *x* is less productive than *y*.
P 2	*a* is a distribution of research & *b* is a possible alternative to *a*.
P 3	*a* leads to an increase of errors or inaction in relevant decisions over *b*.
∴	Distribution of research *a* is less productive than possible alternative *b*.

A third way of arguing that there is a lack of productiveness in *a* is via the efficiency of the distribution: for this type of argument, one does not need to claim that some alternative is less risky, but rather, that the ratio of resources spent to risk reduced is worse in *a* than it is in *b*.

	A1.3: Arguments from efficiency
P 1	If distribution of research *x* leads to a decrease in the ratio of reduction in alethic and practical risks to resources employed over possible alternative *y*, then *x* is less productive than *y*.
P 2	*a* is a distribution of research & *b* is a possible alternative to *a*.
P 3	*a* leads to a decrease in the ratio of reduction in alethic and practical risks to resources employed over *b*.
∴	Distribution of research *a* is less productive than possible alternative *b*.

This allows for imbalances where the alternative is actually not better or even worse in terms of risk reduction, but uses far less of societies' resources than *a*. Note that while in A1.2 it is already specified what constitutes an increase in practical risk—an increase of errors or inaction in relevant decisions—, this is not the case for A1.1. This is because, for the latter argument, there are once again two possible criteria which can support the claim that the crucial premise 3 is satisfied: An increase in alethic risk may either be about the risk of deviating from the truth, or about ignorance.

	A1.1.1: Arguments from alethic error
P 1	If distribution of research *x* leads to an increase in the risk of false beliefs or uncertainty concerning relevant questions over possible alternative *y*, then *x* is less productive than *y*.
P 2	*a* is a distribution of research & *b* is a possible alternative to *a*.
P 3	*a* leads to an increase in the risk of false beliefs or uncertainty concerning relevant questions over *b*.
∴	Distribution of research *a* is less productive than possible alternative *b*.

	A1.1.2: Arguments from ignorance
P 1	If distribution of research *x* leads to an increase in the risk of ignorance about the existence of relevant questions or decision options over alternative *y*, then *x* is less productive than *y*.
P 2	*a* is a distribution of research & *b* is a possible alternative to *a*.
P 3	*a* leads to an increase in the risk of ignorance about the existence of relevant questions or decision options over *b*.
∴	Distribution of research *a* is less productive than possible alternative *b*.

There is one important complication concerning the criteria of alethic and practical risk: As I have said in chapter 5, there are cases where we can only see that there is a problem with alethic risk when we also consider the practical risks involved. This is because the more relevant a research question is, the riskier it is to deviate from the truth in our beliefs about the answer; the relevance of a research question, however, may be influenced by its contribution to practical decisions. In such cases an increase in practical risk may also come with an increase in alethic risk; that is, both A1.1 and A1.2 may be convincing arguments. However—where the increase in alethic risk depends on the practical consequences of error or ignorance—, it would be strange to say that the case is explained by both the criterion of practical risk as well as the criterion of alethic risk. Rather, it should be understood as a problem with practical risk, as it is the risk of error in decision-making that bestows relevance on the research questions at alethic risk.

The other major line of argument support arguments of type A2 by concluding that *a* is unfair. The first way to do is by alluding to the idea of substantive distributive injustice:

	A2.1: Arguments from substantive distributive injustice
P 1	If research distribution x is less productive for members of societal group r than members of societal group t, and this difference cannot be justified, then x is unfair.
P 2	*a* is a distribution of research, *c* is a societal group & *d* is a societal group which is different from *c*.
P 3	*a* is less productive for members of *c* than for members of *d*
P 4	The difference in the productiveness of *a* between members of *c* and *d* cannot be justified.
∴	Distribution of research *a* is unfair.

Note that in the template above, I have expressed distributive risk in terms of a difference in productiveness; here, of course, this difference is not between two alternative distributions of research, but an unequal productiveness where (at least) two particular societal groups—*c* and *d*—are concerned. In chapter 6, I have not endorsed an egalitarian principle of distributive epistemic injustice because there are convincing arguments to the effect that not all cases of unequal productiveness between societal groups should be considered unfair. Therefore, argument template A2.1 includes an additional premise, which states that such differences cannot be justified, and which will have to be supported by more specific moral arguments. I have mentioned the contrast between knowledge as a basic good and knowledge as a private good as one reason for criticizing inequalities that otherwise would not seem problematic.

The other form of epistemic injustice I discussed in chapter 6 is participatory epistemic injustice. The first wrong of participatory injustice is the instrumental wrong: Excluding particular groups from epistemic processes may skew the outcome of the procedure against their favor; i.e., it may be understood as one form of procedural distributive injustice. This idea of instrumental participatory injustice as a form of distributive injustice can be understood as a type of argument to the effect that if there is participatory injustice in the process leading to *a*, *a* will also be substantively distributively unfair:

	A2.1.1: Arguments from instrumental participatory injustice
P 1	If the process leading to distribution of research *x* excluded members of societal group *r* and included members of group *t*, *x* is likely to be less productive for members of group *r* than for members of group *t*.
P 2	*a* is a distribution of research, *c* is a societal group & *d* is a societal group which is different from *c*.
P 3	The process leading to *a* excluded members of *c* and included members of *d*.
P 4	The difference in the productiveness of *a* between members of *c* and *d* cannot be justified.
∴	Therefore: *a* is likely to be less productive for members of *c* than for members of *d*.

Note that the conclusion of A2.1.1 is not identical with the target premise in argument template A.2.1: it just says that a difference in relative productiveness is likely, not that is the case. Therefore, it can only support a probabilistic version of both A2.1 and A2, which, in the end, surmounts to the claim that research distribution *a* likely constitutes a problematic imbalance in the SoR.

What about the last form of injustice discussed in chapter 6, intrinsic participatory epistemic injustice? This criterion should allow us to construct an argument template that directly supports premise 3 in A2:

	A2.2a: Arguments from intrinsic participatory injustice (distribution)
P 1	If the process leading to distribution of research *x* excluded members of societal group *r*, and this exclusion denied members of *r* their deserved epistemic status, *x* is unfair.
P 2	*a* is a distribution of research & *c* is a societal group.
P 3	The process leading to *a* excluded members of *c* and this exclusion denied members of *c* their deserved epistemic status.
∴	Therefore: Distribution of research *a* is unfair.

However, spelling this out in such detail shows that intrinsic participatory injustice only provides a questionable explanation for problematic imbalance in the SoR. Consider premise 1: "If the process leading to distribution of research *x* excluded members of societal group *r*, and this exclusion denied members of r their deserved epistemic status, *x* is unfair". But can we really argue that the distribution is unfair because the process is unfair? In the case of arguments from instrumental participatory injustice (A2.1.1), this inference did not seem problematic. Here, however, the wrongness of being denied one's deserved epistemic status was explicitly supposed to be understood as a wrong intrinsic to the procedure; that is, unlike in A2.1.1, the wrongness of the process cannot be explained in terms of the wrongness of the distribution. Therefore, contrary to the list of criteria in chapter 6, intrinsic participatory injustice does *not* constitute an imbalance in the SoR. Some may object that just because the wrongness of the process is intrinsic, this does not mean it may not also provide cause to believe that the distribution is unfair as well. If this is the case, however, we would still need to argue for why this should be so, that is, why intrinsic participatory injustice should also affect our judgment of the distribution; and this reason cannot be in terms of distributive injustice, for then it would collapse into instrumental participatory injustice. There is some plausibility in thinking that the product of some injustice may be morally deficient in some way even if there is nothing wrong with the product in itself. If we, for instance, believe that the works of art produced by some despicable author—to go for the easy example, think of the paintings of Adolf Hitler—should not be enjoyed, also a SoR produced by unjust science may be thought to be similarly affected. Still, imbalance is defined as a problem with lacunae or overabundance, so this moral tarnish, even if it exists, would not constitute imbalance. An argument from participatory injustice should rather be reconstructed as follows:

	A2.2b: Arguments from intrinsic participatory injustice (process)
P 1	If the process leading to distribution of research x excluded members of societal group r, and this exclusion denied members of r their deserved epistemic status, the process leading to x is unfair.
P 2	a is a distribution of research & c is a societal group.
P 3	The process leading to a excluded members of c and this exclusion denied members of c their deserved epistemic status.
$\therefore$	Therefore: The process leading to a is unfair.

Therefore, intrinsic participatory injustice does not explain imbalance in the SoR. However, as it is part of the original framework, I will still discuss if arguments from intrinsic participatory injustice can be applied to the examples. I also will further discuss the implications of this partial incoherence for the concept of imbalance in subsection 7.2.2.

7.1.3 Application to the Examples

In this section, I will go through the examples in the order that they were introduced in chapter 3 (cf. Table 3.1, p. 86) and discuss, for each of the argument types developed in the previous section, if they are applicable. An overview of the results can be seen in Table 7.1 (p. 224). As is visible there, I distinguish not only two, but three possible outcomes: Firstly, a criterion might be traceable to the literature that first inspired the example. The terminology of the framework was first presented in this dissertation, and thus claims which exactly match the criteria are not to be expected in existing texts. However, the authors of the relevant texts sometimes provide other claims that can explicitly or implicitly support the arguments which represent the criteria. For example, while no one explicitly mentions alethic risk, authors may claim that the SoR in question has problematic consequences for our understanding of the topic, leads to subsequent errors in inferences, or ignorance of important research questions. Practical risk may be suggested, for example, by references to detrimental effects on health decisions or other risks or benefits beyond the effects on knowledge. If the example refers to resources being wasted or mentions an alternative which could save

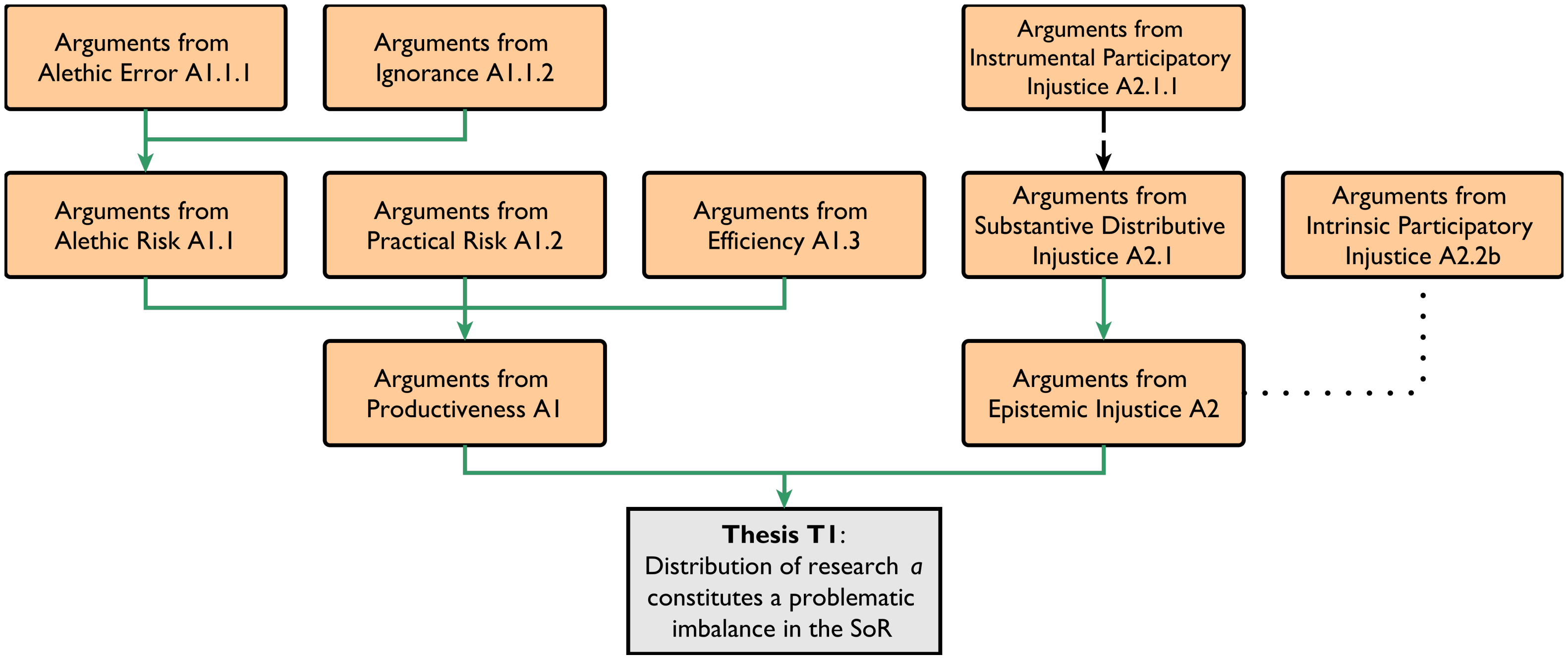

Figure 7.1: Argument map of the operationalized criteria

costs, I count this as a suggestion for an argument from efficiency. As for injustice, when the texts explicitly refer to different groups which stand to benefit or suffer to different degrees, I count this in terms of an explicit support for distributive injustice; if this is connected to a lack of influence on the side of a disadvantaged group, I count it as a reference to instrumental participatory injustice. Lastly, if the sources suggest that this lack of participation also degrades certain members of society, I will claim that the example can successfully be explained in terms of intrinsic participatory injustice. If the literature suggests an explanation in terms of a criterion, it is marked with a "✓". But when a criterion is not explicitly connected to the original example, this does not mean that it is not a possible explanation for a version of the example. Possibility claims, however, are hard to disprove; I do not want to claim that all arguments which are not explicitly excluded are possible reasons to think that an example constitutes imbalance. Instead, I marked an explanation in terms of a criterion as possible ("◊") only if I can provide a tentatively convincing reason for why an aspect of the example may support the respective type of argument.[5] If I could not come up even with such a tentative argument or the example explicitly excludes a criterion as a possible explanation, I marked the application as a failure ("✗").

The problem with regulatory paradigms. The first example refers to regulatory paradigms in the case of the "chlorine sunset"-debate described by Frickel et al. (2010). The authors argued that the regulation of the chlorine-based chemical industry mainly in the Great Lakes region in North America was governed by a "risk"-paradigm, which focused on the "[a]d hoc identification of unsafe chlorine chemicals" and the "[s]ystematic development of chlorine chemicals" while neglecting the "[s]ystematic identification of unsafe chlorine chemicals" and the "[s]ystematic development of nonchlorine alternatives" (all quotes from ibid., 450). This situation we can identify as an instance of the distribution of research *a* which supposedly constitutes imbalance. Considering the possible arguments identified in the previous section, the arguments from productiveness provide a first candidate for a criterion that could justify this judgment. As this type of argument demands, the authors

5. In the previous section, I have argued that when alethic risk becomes apparent only in connection with practical consequences, only practical risk, and not alethic risk should count as the proper explanation. In those cases, alethic risk is also marked with a "◊" rather than a "✓".

do, in fact, consider a possible alternative distribution of research—the "precaution"-paradigm—according to which the focus would be on the "[s]ystematic development of nonchlorine alternatives", but the "[a]d hoc identification of essential and safe chlorine chemicals" would be neglected (Frickel et al. 2010, 450). With that, we have candidates for research distribution *a* and possible alternative *b*.

But how do they fare in relative productiveness, that is, concerning alethic risk, practical risk, and efficiency? There is no explicit evaluation of how many resources either of the paradigms would demand; it is quite possible that developing alternatives to chlorine-based chemicals and a more systematic rather than ad-hoc identification of unsafe chemicals would be more costly than what was actually done. Therefore, there is no clear indication that an argument from efficiency could apply in this case. Also, as I have said above, under both the "risk"-paradigm and the "precaution"-paradigm, some lines of inquiry are in focus and some are neglected; it is not clear from this comparison alone that any paradigm carries more alethic risk in terms of either deviation from the truth or ignorance concerning relevant research questions than the other. What about practical risk, i.e., what are the practical decisions informed by the SoR and what are its goals? The example clearly is about the regulation of unsafe chemicals, and there seem to be two contrary goals in such procedures: firstly, to avoid false negatives in the identification of chemicals and thus avoid "threats to wildlife and humans from persistent, toxic, industrial chlorinated pollutants" (ibid., 449); secondly, to avoid false positives, which would be damaging to the companies whose products would be regulated. Both kinds of risks are relevant to the decision, and again, there would be a trade-off between the two in either of the research distributions considered. However, concerns about risks to the environment and human health could be considered more significant issues than profits, and while regulation should, of course, be fair to the affected companies, its primary goal is the protection of the public. If we take this stance, the dominant "risk"-paradigm carries an increase in practical risk by leading to more errors concerning the protection of health and the environment, which are the more significant errors. Also, this would mean that deviations from the truth in terms of false negatives are more significant and thus relevant than the alethic risk of false positives; as the "risk"-paradigm favors the latter, it would also constitute an increase in alethic error. Similarly, the “risk”-paradigm may make it more likely that relevant issues such as the safety of newly developed chlorine chemicals are not considered, because they have not yet been the target of the ad hoc identification of unsafe chemicals. Thus, both an argument from alethic error as well as an argument from ignorance may be convincingly put forward in this case. However, as I have explained in the previous section, when an increase in alethic risk depends on the practical consequences of the SoR, then only the criterion of practical risk and not the criterion of alethic risk should be considered an explanation suggested by the example.

What about arguments from epistemic injustice? Clearly, the two paradigms are of different value to different groups of people. The dominant “risk”-paradigm is more productive for persons connected to the chlorine industry (social group *d* in argument template 2.1), because this approach is more likely to avoid false positives than false negatives, and also suggests the active development of new chlorine-based alternatives. The “precaution” paradigm, however, is more beneficial for people who stand to be harmed by the environmental hazards that could be avoided by a systematic identification of unsafe chlorine chemicals (social group *c* in the argument template). Therefore, the dominant “risk-paradigm”-could be considered unfair for those who are affected by the environmental hazards.[6] However, it could be argued that also the “precaution”-paradigm is unfair, only to the other group. This would suggest a stalemate, where no distribution of research is better than the other in terms of fairness. This then could be used to argue that there is a justification of the uneven distribution of productiveness—which would, as per the fourth premise, disqualify an argument from substantive distributive injustice. However, we have better reasons to believe that the “risk”-paradigm, and not the “precaution”-paradigm is unfair: In chapter 6, I have said that when we weigh basic interests—and human health surely is one of those—against private interests—such as the profits of companies in the chlorine industry—the former should be assigned greater weight. Therefore, there is actually no justification for the difference in productiveness, and we can argue that

6. Both the role of the “powerful chemical industry” and “extensive citizen activism” (Frickel et al. 2010, 449) are explicitly mentioned by the source text, substantiating the idea that justice might provide an additional argument for why the case constitutes a problem.

the "risk"-paradigm violates the criterion of substantive distributive injustice.

In the description of the case, it is not made clear if there also is a participatory injustice, perhaps, in the sense that industry representatives were able to influence the scientific process and citizens or non-governmental organizations were not. Instrumental participatory injustice, therefore, does not clearly provide a reason for why this case should be considered an imbalance. As for the intrinsic wrong of participatory injustice, even if there is a problem with exclusion in this case, it would probably not be considered denying someone their deserved epistemic status, and thus this criterion cannot be applied to this example.

To summarize, arguments from practical risk (A1.2) and alethic risk (A1.1, A1.1.1, A1.1.2) can be successfully applied to the case of regulatory paradigms. However, as the alethic risk results from the practical consequences of the research, only the criterion of practical risk provides a proper explanation. With A2.1, we could alternatively conceive of it as a problem with distributive injustice. Furthermore, it could be possible to provide an explanation in terms of or instrumental participatory injustice (A2.1.1). However, this is not suggested by the source material. Lastly, I could provide no support for an argument from efficiency (A1.3), and A2.2 is clearly not applicable in this case.

The problem with unpatentable research. The second example was taken from Brown (2008), who argues that in medical research, lines of inquiry which may lead to patentable products are in focus, and alternative approaches are neglected:

> Imagine two ways of approaching a health problem. One way involves the development of a new drug. The other way focuses on diet, exercise, or environmental factors. The second could well be a far superior treatment, both cheaper and more beneficial. But obviously it will not be funded by corporate sponsors, since there is not a penny to be made from the unpatentable research results. It should be just as obvious that a source of funding that does not have a stake in the outcome but simply wants to know how best to treat a human ailment would, in principle, fund either or both approaches, caring only to discover which is superior. (ibid., 197)

This neglect of unpatentable solutions in favor of drug development is our distribution of research, *a*. But which of the criteria can suggest an explanation for why it is a problem? The mention of funding sources and stakes seems to suggest a problem with participatory injustice: members of one group—sponsors with interest in the profits generated by patentable solutions—can influence the research process, while others—such as patients—have less say. It is not quite clear if Brown criticizes the SoR as substantively more beneficial to the former than for the other; however, as per the argument from instrumental participatory injustice, it could be argued that caring only for which of the ways is the superior treatment would make it more likely that the SoR is equally productive for persons with commercial interests and patients alike. Once again, even if this would skew the SoR in favor of being productive for human health over profits, the former is a basic, and the latter only a private interest. Thus, the criterion of instrumental participatory injustice provides a first explanation for why the example of unpatentable research may *likely* be problematic. Intrinsic participatory injustice, however, does not seem to be applicable: it is not clear why not being able to influence the funding process would constitute an insult to anyone in their epistemic status.

But what about productiveness? If it is *either* research on patentable solutions *or* research on the alternatives, it is not clear that, generally, any of the two would be preferable in terms of alethic or practical risk: while, as Brown says, the latter "could well be a far superior treatment" (ibid., 197) this is by no means guaranteed. Still, funding decisions are generally not made in empty space; usually, we may have some reason—perhaps, due to the existing body of knowledge and the research proposal—that one project is more promising or significant than the other. In those cases, then, not judging the alternative projects by their prospective quality, but by their prospective market value would likely increase both alethic risk—where indicators for the relative contribution to our understanding or knowledge are ignored—and practical risks—where considerations of profit trump estimates of the projects' contribution to health decisions. How convincing such an argument is, thus hinges on our ability to judge the relative productiveness of different lines of inquiry in advance.

This is not so, however, if it would be a possible alternative (*b*) to follow up on both, perhaps by slightly reducing funding for patentable solutions instead of giving up on it altogether. For this alternative, Brown provides an argument to the effect that *b* actually is less alethically risky than *a*. As I have mentioned in my discussion of the concept of diversity in subsection 5.2.1, funding competing research projects with the same goal may—because of the benefit of competition, and also because of diminishing marginal utility—increase the likelihood that we get closer to the truth. Brown clearly considers diversity beneficial "for the sake of improving the growth of knowledge" (Brown 2001, 187), i.e., as a general epistemic principle. In argument templates A1.1.1 and A1.1.2, this argument via diversity could support the third premise. Thus we can consider a more diversified alternative to the SoR in the case of unpatentable research to be beneficial in terms of alethic risk. Not only in terms of alethic error, but also in terms of ignorance, a diversified SoR may be considered superior: even a cursory investigation of alternative approaches may indicate new research questions and decision options connected to the relative benefits and disadvantages of both drug-based approaches and the unpatentable alternatives. Furthermore, as medical research is clearly connected to practical decisions about human health, if we accept Brown's argument, also an argument from practical risk can provide an explanation for why the case of unpatentable research is problematic. While there is no clear mention of considerations of efficiency, it is possible to construct an argument from diminishing marginal utility, claiming that a unilateral focus on drug-based interventions is most probably not the most efficient use of research resources.

To summarize, unpatentable research can be explained as a problem with instrumental participatory injustice (A2.1.1b)—which would only amount to the claim that it likely is a problem—but also, perhaps closer to Brown's original description, as a problem with a lack of diversity, leading to increase in alethic (A1.1, A1.1.1, A1.1.2) and practical (A1.2) risk. Arguments from efficiency (A1.3) and substantive distributive injustice (A2.1) may be possible but are not suggested by the source. There was no support for an argument from intrinsic participatory injustice (A2.2).

The problem with neglected diseases. The example of neglected diseases was prominently featured in chapter 6: There, I argued that we can explain this case both as a problem with productiveness and as a problem with epistemic injustice. The idea is that, once again in the area of medical research, diseases which cause great suffering in less well-off parts of the world receive far less attention than well-explored and perhaps less dramatic health problems afflicting people in wealthy nations. The focus of authors such as Flory and Kitcher (2004) and Reiss and Kitcher (2009) is clearly on the practical risk caused by neglected diseases, arguing that the suffering that could be avoided by bringing them into focus would be far greater then what is achieved by the current research agenda. We can reconstruct this both as an argument from practical risk (A1.2)—focusing on the risk reduced alone—and as an argument from efficiency (A1.3), as the authors explicitly lament that the ratio of suffering reduced to research dollars spent on diseases is problematically low under the current system. The problem of practical risk could also allow for arguments from alethic error and ignorance—perhaps, again, in terms of the diminishing marginal utility of spending more money on problems that are already relatively well-understood compared with diseases which have hardly been under scientific scrutiny. But as this risk stems from the practical consequences of erring or not knowing about relevant research questions, I do not consider it a main explanation.

However, the example is clearly not just about a lack of productiveness: the problem is regularly framed as being about "diseases that afflict poor nations" (Flory and Kitcher 2004, 46), which suggests that the case is also about *who* is served by science. Clearly, the case can thus be explained in terms of substantive distributive justice: members of less affluent societies are at much higher risk from neglected diseases than members of affluent nations. One may object that the latter also pay for most of the research, and thus that the difference in productiveness is justified. Depending on one's ethical stance, this may be a valid objection. If, however, one accepts ideas such as the Rawlsian difference principle, we would have to label the inequality unfair: according to the second principle of justice, "[s]ocial and economic inequalities are to be arranged so that they are [...] to the greatest benefit of the least advantaged [...]"(Rawls 1971, 83). As at least some of the people in the poorest parts of the worth both suffer from neglected diseases and should be counted among the least advantaged, this distribution of research violates the principle.[7]

7. Note that this argument also requires that the difference principle is not just applied to a particular society—to the citizens of one nation, for example—but to all humans. This, however, is not a feature of an argument from distributive justice alone: all possible explanations of the example of neglected

The problem of neglected diseases is also explicitly described as a result of funding mechanisms:

> Research dollars come almost entirely from the wealthy parts of the world, and the suffering from malaria, tuberculosis, and a large number of other infectious agents happens elsewhere. (Flory and Kitcher 2004, 40)

While it would thus be possible to construct an argument from instrumental participatory injustice (A2.1.1), this would be superfluous as we already have a stronger argument from substantive distributive injustice (A2.1). Also in the example of neglected diseases, it is not clear that the problem is connected to anyone being denied their deserved epistemic status, so an argument from intrinsic participatory injustice (A.2.2) would not be convincing.

To summarize, neglected diseases can primarily be explained as a problem with practical risk (A1.2), efficiency (A1.3), and substantive distributive injustice (A2.1). The substantive claim trumps a possible argument from instrumental participatory injustice (A2.1.1). Indirectly, we can also conceive of neglected diseases in terms of alethic risk (A1.1, A1.1.1, A1.1.2). The example does not suggest an explanation in terms of intrinsic participatory injustice (A2.2).

The problem with ignored genes. The example of ignored genes refers to another criticism of biomedical research according to which research has focused on a relatively small number of proteins involved in human genetics, and equally relevant targets of investigation have been neglected. Edwards et al. (2011) mention two significant aims, which are negatively affected by this distribution of research: to "transform our understanding of human biology and disease, and provide new targets for drug discovery". With most of research focusing on the same few targets, we can expect diminishing returns and thus an increase in alethic risk concerning "our understanding of human biology and disease". This indicates that our beliefs about these topics may be more likely to be wrong, but also, as the talk about "new targets" suggests, that we are more likely to remain ignorant about important research questions and options for treating disease. Especially the latter may also be interpreted

diseases need to presume this perspective for them to be convincing.

as a practical risk in terms of making suboptimal health decisions. We might suspect that—because of diminishing returns—it would be more efficient to invest in a diversified agenda. However, Edwards et al. (2011) emphasize that research focusing on other proteins would also involve substantial costs, especially because it also requires researchers to develop new methods:

> For example, the level of funding needed to develop even one chemical probe is enormous. Although it is only a fraction of the US$100 billion spent on biomedical research each year—about several million dollars—it is huge compared with the amount customarily allocated to an individual scientist. (ibid.)

Thus there is no clear case for an argument from efficiency.

As for explanations in terms of injustice, one could argue that the current system mostly favors scientists and their career plans, and neglects what is best for the patients, and thus there is an element of distributive or instrumental participatory injustice (A2.1 or A2.1.1); this interpretation, however, is not directly implied in the example. Once again, there is no indication of anyone being denied their deserved epistemic status, excluding an argument from intrinsic participatory injustice.

To summarize, the example of ignored genes can be explained both as a problem with alethic (A1.1, A1.1.1, A1.1.2) and with practical risk (A1.2). Arguments from distributive injustice (A2.1 or A2.1.1) may be possible but are not suggested by the source material. There is no support for an argument from efficiency (A1.3) or intrinsic participatory injustice (A.2.2).

The problem with the evidential standards of EBM. The example of the evidential standards of EBM suggests that evidence hierarchies supported by the "evidence-based medicine"-movement may lead to imbalance: these hierarchies value some forms of evidence—such as randomized controlled trials, meta-analysis, and systematic reviews—generally much higher than others; based on non-randomized studies or individual patient histories, for example. This may lead to those latter types of research being excluded from the SoR (cf. Stegenga 2018, 73, footnote 5).

Jacob Stegenga (2011) points out that this would constitute an intrinsic epistemic problem because, among other reasons, it violates the principle of total evidence.

About the quality of the method of meta-analysis—a method which, as I have argued in chapter 2, can be considered a reconstruction of the SoR—, he writes:

> [N]ot only is evidence from non-randomized studies not to be amalgamated with evidence from RCTs, but neither is evidence from pathophysiological knowledge, background considerations of underlying mechanisms, animal experiments, and results from mathematical models. Such a practice could limit the external validity of a meta-analysis[...]. Moreover, as discussed above, this practice violates a principle of total evidence, which comes with possibly significant epistemic risk: neglecting other kinds of evidence risks making an uninformed judgment (or, the base-rate fallacy) on a hypothesis. (Stegenga 2011, 501)

This idea of epistemic risk connected to the judgments on subsequent hypotheses is very close to the criterion of alethic risk, and, more specifically, alethic error. In another paper critical of the EBM movement, Rachel Ankeny (2014) emphasizes that case-based research also provides research questions that can subsequently be explored with other methods:

> As such, cases often are essential first steps within contemporary research practices that provide working candidates for further tests of causal relations using more conventional methodologies such as RCTs or other experimental methodologies, or retrospective case controls. (ibid., 1001)

This suggests that an exclusive focus on methods ranked highest in some evidence hierarchies may cause such "working candidates" to be overlooked and thus support an argument from ignorance.

Also the example of the evidential standards of EBM is mostly about biomedical research. In this context, there are concerns not only about the implications for knowledge in itself, but also for decisions about, for example, drug approval. Therefore, also arguments from practical risk provide a possible explanation for why it may be problematic.

For a convincing argument from productiveness, we also need to provide a possible alternative to a distribution of research based on an exclusive EBM hierarchy.

While it would be easy to suggest that one should just use all the available evidence instead of excluding some methods, some are clearly superior to others in at least some respects; to arrive at a summative judgment, different sources of evidence will therefore have to be weighed against each other. Rather than arguing for a specific alternative distribution, however, it may be more promising to caution against interpreting evidence hierarchies to strictly, discuss the relevance of the evidence case by case, and not exclude any method of some value in principle. Though he is also critical of these methods—mainly because of bad inter-rater reliability—Stegenga (2018) suggests so-called quality assessment tools for this job (cf. ibid., chapter 7). Given this difficulty of how to include other types of evidence, it is also not clear how resource-intensive alternative approaches would be, and I cannot suggest a convincing argument from efficiency.

What about epistemic injustice? While the authors mentioned above do not explicitly consider a connection between EBM standards and issues of fairness, already in subsection 6.2.2 I have mentioned the idea that evidence hierarchies may exclude certain people from the scientific process: if case-based research were shunned, also the individual experiences of patients may be ignored, constituting a form of participatory injustice. Kidd and Carel (2017) see it as an advantage of narrative medicine that this alternative approach teaches "[m]edical students, trainees, and practitioners [...] how to listen to patient narratives and how to utilise those narratives in clinical practice" (ibid., 344). Lacking such a role for patient testimony, both forms of participatory injustice may provide a reason to criticize evidence hierarchies. Firstly, it deprives patients in their "*capacity to provide information* that can meaningfully contribute to the epistemic task" (ibid., 339), which may be understood in terms of an argument from instrumental participatory injustice; secondly, "in cases of testimonial injustice, patients are perceived as 'somewhere between an epistemic subject and object'"(ibid., 343), indicating that they are deprived of their deserved epistemic status, which can support an argument from intrinsic participatory injustice.

To summarize, if the evidence hierarchies of EBM lead to an outright exclusion of certain types of evidence, the resulting research distribution could be understood both in terms of alethic (A1.1, A1.1.1, A1.1.2) and practical (A1.2) risk. Together with ideas about healthcare discussed in the debate about epistemic injustice, we may also try to explain the problem with this case as an issue with both instrumental (A2.1.1) and intrinsic (A2.2) participatory injustice. While A.2.1.1 suggests that distributive injustice is a likely explanation, there is no independent argument for substantive distributive injustice (A2.1). There is no support for an argument from efficiency (1.3).

The problem with air-monitoring standards. The example of air-monitoring standards refers to an alleged problem with the role of standards concerning the regulation of air quality, where these may act in a "boundary-policing function" (Frickel et al. 2010, 455): evidence collected by activists and citizens which reports high concentrations of toxins in the air—perhaps caused by nearby factories—may be excluded by the standards, because this so-called "bucket data" does not conform to the type of data required:

> Specifically, ambient air standards are typically expressed as averages over a period of hours, days, or years. Bucket data, in contrast, characterizes average chemical concentrations over a period of minutes. (ibid.)

Thus, the standards may mark these contributions of laypeople as irrelevant to the regulatory procedures, excluding them from the SoR. Bucket data may help scientists and officials to deal with the "dearth of scientific research that could illuminate, in ways credible to residents, the effects of industrial emissions on community health"(ibid., 454) and thus "address questions of pressing concern to community members but hitherto ignored by experts" (ibid., 455). This can be thought of as an improvement in terms of alethic risk. Where the use of bucket data can point towards cases of air pollution that the experts have missed, this avoids false negatives, reducing alethic error. The authors also mention further positive consequences of including the activists' data:

> To the extent that bucket monitoring has resulted in increased enforcement activity by regulators [...] or additional ambient air monitoring by industrial facilities, the additional monitoring has been undertaken to confirm activists' results, track the causes of the chemical emissions, and fix what are assumed to be isolated malfunctions [...]. (ibid., 456)

This quote suggests more than just benefits in terms of alethic risk: when chemical emissions would not even have been detected, important research questions would not have been raised, implying a problem with ignorance. Being able to fix malfunctions, and, of course, the context of regulating air pollution in general, suggests an advantage in terms of practical risk.[8] With the authors' mention of bucket data as an "alternative, more cost-effective method" (ibid., 456), there is also an explicit suggestion for an argument from efficiency.

But is the example of air monitoring standards also a case of epistemic injustice? The authors clearly suggest that the inclusion of the bucket data can help detect problems that could cause health problems for the local communities. On the other hand, it might also be problematic for those who are invested in the air-polluting industry; members of that group may suffer as a consequence of increased regulation. As the authors put it, the exclusive standards can "give industrial facilities and environmental agencies a ready-made way to dismiss bucket data" (ibid., 456), suggesting the practice may be considered a form of participatory injustice: because citizens lack the opportunity to contribute to the research and thus to the regulatory process, a SoR policed by exclusive air-monitoring standards may be less productive for them than for those connected to the industry, indicating instrumental participatory injustice. It may also be argued that this exclusion denies members of local communities their proper epistemic status as local experts; this, however, is not explicitly suggested by the text. Also, besides the argument from instrumental participatory injustice, there is no explicit claim supporting an argument from substantive distributive justice.

To summarize, the example of air-monitoring standards could be taken as a problem with all criteria connected to productiveness (A1.1, A1.1.1, A1.1.2, A1.2, A1.3). Also, it may be seen as a case of instrumental (A2.1.1) and perhaps, though not explicitly, as a case of intrinsic participatory injustice (A2.2); though possible, there is no case for an explanation in terms of distributive injustice (A2.1) which is

8. Some, however, also explicitly dispute this claim: While activists apparently decry this lack of influence of the bucket data, the authors also mention that according to some experts "only the average concentrations of regulated chemicals can be meaningfully compared to the standards and thus contribute to determining whether air-pollution might pose a threat to human health."(Frickel et al. 2010, 456).

independent from A2.1.1.

The problem with publication bias. The example of publication bias refers to a problem where—irrespective of quality—prospective scientific publications reporting negative results are less likely to be published than those reporting positive results. The meaning of "positive" and "negative" is somewhat vague in this context; to be more precise, publication bias is often analyzed in terms of statistically significant vs. non-significant effects, but other contrasts may be applicable as well.

Why is publication bias a problem? We can easily see that in terms of deviation from the truth, a lacuna of negative reports concerning some hypothesis or theory or method may systematically skew the beliefs of recipients of the SoR towards erroneously believing that the hypothesis is true, while they at least should be more doubtful. For example, consider this description of the effects of publication bias in basic research:

> If positive results from basic research are more likely to be published than negative results, results of published studies of basic research will represent an overestimation of potential treatment effects. (Song et al. 2010, 39)

Similar to the last example, we can thus support an argument from alethic error with some version of a principle of total evidence—which is violated by publication bias. If there exists no other study focusing on the same research topics besides the ones that remain unpublished due to publication bias, or the existing research does not touch on certain aspects which the unpublished texts would have, also an argument from ignorance of research questions or decision options may be applicable, although this is not usually an aspect prominent in the discussion. The remaining aspects of productiveness, however, are explicitly mentioned: " The most important consequences of publication bias include avoidable suffering of patients and waste of limited resources." (ibid., 40) As this quote suggests, when publication bias affects areas of research with consequences for practical decisions, it likely also increases practical risk. Also, publication bias wastes resources: we may still interpret this as a kind of practical risk, where health decisions based on a biased SoR may lead to the usage of expensive but ineffective treatments; but already when considering the production of the research itself, if some study has been conducted but cannot be published although it is of high quality, it is certainly also wasteful in terms of an argument from efficiency. With arguments from productiveness, we also have to establish that there are possible alternatives to publication bias, which would be more productive. There are two lines of thought: firstly, we may try to improve the parts of the scientific process that cause publication bias and aim for a more balanced SoR while keeping the overall amount of publications stable. For example, to counteract problems with scientific journals, "editorial policy needs to be changed to allow the acceptance for publication of clinical trials that are based only on methodological criteria and not on the impact of their findings" (Song et al. 2000, 40); also preregistration is frequently mentioned (Song et al. 2000; Song et al. 2010). Secondly, one could also simply try to publish more studies overall, including those which were inaccessible before. While this may have been impossible in earlier times, with the advent of electronic publications, space-restrictions which would have excluded this possibility theoretically have fallen (cf. Song et al. 2010, 52-53). It should be noted that publication bias also constitutes a communication problem. That is, at least some of its negative effects in terms of productiveness could be avoided if it can be detected and communicated to the recipients of the SoR.

Can the problem with publication bias also be understood as a kind of injustice? There are various possible causes of publication bias, some of which may also indicate that someone stands to profit more from the resulting SoR than others, which could allow for arguments from substantive distributive as well as instrumental participatory injustice. However, the texts about publication bias discussed here do not provide any such a rationale. Even if there are participation deficits concerning the publication process, however, these would not seem to infringe upon anyone's deserved epistemic status.

To summarize, publication bias can be explained as a problem with productiveness in terms of alethic error (A1.1.1), practical risk (A1.2), and efficiency (A1.3). Arguments from ignorance (A1.2) and distributive justice (A2.1, A2.1.1) may also be applicable but are not explicitly suggested. An argument from intrinsic participatory injustice (A2.2) would not be convincing.

The problem with suppression. The example of suppression is about cases where parts of publications—or, more relevant for a discussion of imbalance, also whole studies—go unreported because they are actively kept secret by an interested party—often the sponsors of the studies, or perhaps the authors. While this may sometimes also result in publication bias, what goes unreported here goes beyond negative results. Everything that could be damaging to some party with power over what is reported might be suppressed. Sometimes, that may mean negative results—for example when the results of a study show that the product of a sponsor does not have a statistically significant effect. But suppression may just as well affect positive results, for example, when they indicate that the sponsors' product has adverse effects.

Like with publication bias, also in the case suppression, if not all the existing evidence can be considered, this may cause subsequent inferences to diverge from the truth, leading to an increase in both alethic risk and practical risk.[9] With the risk of "unnecessary duplication" (Chan et al. 2014), authors explicitly mention an effect of suppression that could support an argument from efficiency—and also more generally, where high quality research is held back, this is certainly wasteful. In cases of suppression rather than publication bias, it is also more plausible to argue that it increases the risk of ignorance: here, also positive results attesting that, for example, some drug has some unforeseen adverse effects which should be investigated or considered when making health decisions could go unnoticed. Stegenga (2018) reports multiple examples for trials reporting negative side-effects being held back (cf. ibid., 146-149), supporting an argument from ignorance.

In the case of suppression, it is also clearer how arguments from justice could be justified: If sponsors of scientific studies—or perhaps authors with vested interests—have sole control of what is reported, the research process excludes those affected by the research. In consequence, only the suppressors' interests would be served by the SoR. Still, suppression is not usually described as a matter of justice. Where certain actors actively try to keep valuable information secret because it would be damaging to them, we might want to describe this as a case of malevolent deception, rather than just distributive injustice. In those cases, not only can the resulting SoR be considered out of balance, but—because of a lack of benevolence—it also directly constitutes

9. See Chan et al. 2014, 2 for a list of examples of selective reporting and the resulting damage.

a problem for the trustworthiness of science. In any case, intrinsic participatory injustice cannot provide a convincing argument, as even if we frame this in terms of a lack of participation, no one is clearly deprived of their deserved epistemic status.

To summarize, suppression can be explained using each and any of the criteria connected to productiveness (A1.1, A1.1.1. A1.1.2, A1.2, A1.3). While, compared with publication bias, there is greater cause to suspect that suppression may also be distributively unfair (A2.1, A2.1.1), this is not clearly the best explanation for what is wrong about the actions of the suppressors. An explanation in terms of intrinsic participatory injustice (A2.2) would not be convincing.

The problem with multiple trials with predictable outcomes. Sergio Sismondo (2008) provided the example of multiple trials with predictable outcomes. He suggests that there is an overabundance of research on so-called "blockbuster drugs"—but what does overabundance mean in this context, i.e., why is it *too much*? We might say that—when compared with alternative distributions of research that include more studies on, for example, "drugs of similar age but with small patient populations"—multiple trials with predictable outcomes lead to an increase in alethic or practical risk: Because of the prevalence of blockbuster drugs in the literature, doctors and patients might wrongly assume that there are no alternatives to these drugs—which would support an argument from ignorance—, or underestimate how beneficial alternatives are—an increase in alethic error—, thus leading to suboptimal health decisions—i.e., to practical risk. However, Sismondo's example explicitly highlights the predictable outcomes of these studies, which are "designed to test an already-studied drug in a way known to be effective, on a population for which it is known to be effective". This mainly suggests a problem with diminishing marginal utility, and a lack of efficiency where there are many such trials, wasting research resources.

One could also once again weigh the benefits for the financially interests sponsors against the benefits for patients, and argue that this constitutes a substantive distributive injustice, but also here, this is not explicitly suggested by the original example. Participatory injustice is even less convincing, as this is mostly about privately funded research, and it is unclear why patients or other groups should be included in decisions about what private companies do or do not sponsor.

To summarize, the example of multiple trials with predictable outcomes is mainly framed in terms of a lack of efficiency (A1.3). The example could theoretically also be explained in terms of alethic risk (A1.1.1, A1.1.2), practical risk (A1.2), or substantive distributive injustice (A2.1). Arguments from participatory injustice (A2.1.1, A2.2) are not convincing for an example concerning agenda setting in private research.

The problem with the "pause" in global warming. The example concerning the alleged "pause" in global warming is about another form of overabundance, where authors like Lewandowsky, Risbey, and Oreskes (2016) argue that there was too much research on certain temperature fluctuations. Critics of climate change, according to the authors, had referred to these fluctuations to argue that "global warming—measured by global mean surface temperature (GMST)—has 'stalled,' 'stopped,' 'paused,' or entered a 'hiatus'"(ibid., 723).

In contrast to the example of multiple trials with predictable outcomes, there are no explicit references to efficiency as a possible explanation. According to the authors, it is at least not clear that this overabundance of research on the "pause" has wasted research resources, or, in fact, caused any direct problem with productiveness[10]:

> The body of work on fluctuations in warming rate has clearly contributed to our understanding of decadal variations in climate. [...] Research on the pause has thus ultimately reaffirmed the overall reliability of climate models for projecting temperature trends. (ibid., 729-730)

Still, the authors argue that the overabundance of research referring to the "pause" is a problem. Recipients both within and beyond science may take the overabundance of studies concerning the topic as an indicator that there was something to be explained:

> However, by accepting the framing of a recent fluctuation as a pause or hiatus, that research has, ironically and unwittingly, entrenched the notion of a pause

10. If we accept the argument that the focus on the pause—via a communication problem—has caused epistemic and practical risk after all, also an argument from efficiency becomes viable, as there is no cause to believe that alternative lines of inquiry would have been more costly.

> (with all the connotations of that term) in the literature as well as in the public's mind. [...] Accepting contrarian linguistic frames helps maintain the fiction that the science is still too uncertain to form a reliable basis for public policy. (ibid., 730-731)

We can think of this as a communication problem, where recipients within science and society misinterpret the research to mean that climate science is less reliable than it is. If the research on the "pause" thus leads to an increase in uncertainty about, for example, the reality of anthropogenic global warming, it constitutes a problem of alethic error when compared to alternative distributions where other topics in climate science would have been explored; if it leads to undue hesitation or a lack of support of much-needed policy decisions, it also constitutes a case of increases practical risk. There is no indication that the focus on the "pause" led to increased ignorance about important research questions or decision options.

In Lewandowsky et al. (2015), the authors suggest that the overabundance of research on the alleged pause may be the result of an effect called seepage, where the claims of contrarians and climate change skeptics had an undue influence on the research agenda. While we may see this as a problematic influence on the scientific process, it is unclear how this could be construed as a case of participatory injustice, as no particular group was excluded from epistemic activities. Still, we may see it as a distributive injustice because the resulting research may be of more relevance to the questions of the skeptics—and by seeding doubt may have also furthered their practical goals—, than to the interests of people who are concerned about the effects of anthropogenic global warming. However, no interpretation in terms of justice is explicitly suggested by the discussion in the original texts.

To summarize, the example of the alleged pause in climate change can best be explained as a problem with alethic error (1.1.1) or practical risk (1.2). There is no cause to suspect that the focus on the "pause" has led to an increase in ignorance (A1.1.2). An argument from efficiency (A1.3) may be possible, but is not clearly suggested. It would also be possible to argue that it is a case of substantive distributive injustice (A2.1), but explanations in terms of participatory injustice (A2.1.1, A2.2) would be far fetched.

	A1.1.1	A1.1.2	A1.2	A1.3	A2.1	A2.1.1	A2.2
Regulatory paradigms	◊	◊	✓	✗	◊	◊	✗
Unpatentable research	✓	✓	✓	◊	✗	✓	✗
Neglected diseases	◊	◊	✓	✓	✓	◊	✗
Ignored genes	✓	✓	✓	✗	◊	◊	✗
Evidential standards of EBM	✓	✓	✓	✗	◊	✓	✓
Air-monitoring standards	✓	✓	✓	✓	◊	✓	◊
Publication bias	✓	◊	✓	✓	◊	◊	✗
Suppression	✓	✓	✓	✓	◊	◊	✗
Multiple trials with predictable outcomes	◊	◊	◊	✓	◊	✗	✗
“Pause” in global warming	✓	✗	✓	◊	◊	✗	✗

Table 7.1: Overview: criteria & examples

7.2 Discussion

7.2.1 Evaluation of the Results

The discussion of productiveness and justice in chapter 5 and chapter 6 was supposed to provide a comprehensive and coherent framework that can be used to explain what is wrong in the examples of imbalance presented in chapter 3. The criteria are very successful in terms of comprehensiveness: As Table 7.1 (p. 224) shows, for each of the examples, more than one of the criteria provides a possible explanation; moreover, the texts from which the examples where taken can be used to directly support at least one argument based on the criteria in each case. It could be objected that, while I can always explain at least one of the intuitions presented by the original authors of the examples, I cannot always account for every aspect mentioned in the source material. One such aspect might be the wrong of fraud or deception, which is evident in the example of suppression. However, fraud and deception in themselves are no intrinsic aspects of the research distribution; a SoR resulting from these wrongs only constitutes imbalance if it is also unproductive or unjust. As such, one could include fraud and deception as procedural criteria, similarly to A2.2.1 or a lack of diversity that is detrimental to productiveness. Alternatively, we can describe fraud and deception as problems with benevolence, and thus with the trustworthiness of science. While there are thus perhaps some aspects of the cases that my account does not cover, it is always difficult to prove that no possible interpretation has been overlooked. As far as problematic aspects of the distributions themselves go, there are no clear gaps left unaccounted for. Moreover, the framework proved very fruitful in terms of suggesting additional possible explanations that were not implied by the authors of the original examples.

At the same time, the criteria provided are also coherent. Already in subsection 7.1.1, I explained that independent from the examples themselves, there are relations between the criteria of productiveness, epistemic justice, and trust on a conceptual level. But coherence also became evident during the application of the criteria to the examples. There were no clear contradictions between any judgment suggested by the framework and the original author's claims or any obvious intuitive judgments of the cases. In terms of inferential connections between the criteria, the argument map in Figure 7.1 shows how many of the criteria and sub-criteria build on each other and can be used to support a common conclusion. While there are two distinct lines of argument portrayed in the map, the analysis of the examples does not suggest that problems with productiveness and justice are unrelated: in all of the cases, both arguments from productiveness and justice were deemed at least possible, and in multiple examples, both where suggested by the source texts. Thus, the framework does not just artificially connect two or more wholly independent issues; instead, the concept of imbalance presented in this dissertation at the very least describes problems with research distributions as a shared nexus of problems with the overall usefulness of the SoR and its fairness. I will explore the implications of this in subsection 7.2.2.

Interestingly, practical consequences of the SoR seem to be at the heart of most cases of imbalance: All but one of the examples suggested practical risk as a consequence of imbalance. Only in the case of multiple trials with predictable outcomes practical risk is not directly suggested as an explanation. But even here one could argue that the claim that many of the trials "are designed more to 'familiarize ' physicians and patients of products than to produce novel knowledge" (Sismondo 2008, 3) hints at negative effects on health decisions; this connection, however, was too shallow for me to count this as a case where the original example implies practical risk. The prevalence of practical risk is not surprising; exclusive problems with alethic risk would suggest that there is a total disconnect between the scientific knowledge and its practical applications; and where only curiosity is at stake, criticisms of science may not be as urgent. There could also be some selection bias involved: already in chapter 3, I remarked that a majority of the examples come from areas of science highly significant for societal problems—with a heavy emphasis on biomedical research. It would be interesting to see if practical risk is similarly prevalent in fields traditionally less application-oriented—theoretical physics, perhaps. However, examples of imbalance are at least much harder to find in those areas. Perhaps also this is not entirely surprising: in the introduction of this dissertation, I have argued that the perspective from the level of the SoR is especially relevant where we have very complicated societal issues. The SoR is of importance here because many different research projects could provide answers for the relevant research questions. But of course, we worry about these complicated issues because also the practical stakes are high. While this may explain why practical, rather than alethic risk is in the focus of the examples, one might expect stronger connections to matters of justice. Distributive injustice was considered a possible explanation in basically all of the cases. It should be noted, however, that this required a rather broad interpretation of differences in productivity between groups; I often refer to different benefits for different groups, where it not always clear that they can be expressed in terms of negative impacts on decision making, for example.

The argument map also suggests, however, that there is one criterion that *is* isolated from the others. The application to the examples further reinforced this impression of an anomaly, as intrinsic participatory injustice was rarely applicable. The examples were selected as problems with the SoR, or more specifically, the research distribution. Thus, a disconnect with an intrinsic procedural criterion does not come as a surprise. This does not mean that that intrinsic wrongs of the procedure are entirely distinct from imbalance in the SoR; in the discussion about the epistemic status of patients, for example, I suggest a connection between imbalance and intrinsic participatory epistemic injustice.

In summary, the conceptual framework presented in the last chapters does very well both in terms of comprehensiveness, as well as—with the exclusion of intrinsic participatory injustice—coherence. Beyond that, the framework can also suggest further possible explanations that could apply to these cases, and merit additional exploration of the underlying problems.

7.2.2 Imbalance and Higher-Order Concepts

In section 3.2, I discussed several options for possible concepts that could replace the term imbalance: mainly, these were "freedom" or "independence", "conflicts of interest", "bias", and "objectivity". I argued that objectivity may be the most promising candidate, but that this concept has so many interpretations that it cannot straightforwardly apply. Instead, I proceeded by examining trustworthiness—which some claim is related to objectivity—and its constitutive elements. This was not only supposed to provide a more systematic account of what is wrong in cases of imbalance—which I have done in the previous sections—but also contribute to our understanding of the concepts mentioned above.

How does my discussion of imbalance contribute to our understanding of scientific independence and freedom? There are multiple connections between these concepts: Firstly, I have argued that granting scientific freedom can be understood as a form of public trust in science that goes beyond of what is demanded by a difference in expertise. I argued that by learning what reasons there might be for granting scientific freedom, we might also expose some requirements for trust that are especially relevant in the relation of science and the public. The resulting criteria—productiveness and epistemic injustice—, are the same that can be used to explain our intuitions about the examples of imbalance. What about independence? Scientific independence or autonomy, as the absence of external restrictions on decisions made in the process of science, can be considered the consequence of scientific freedom. If we believe the arguments for scientific freedom, then this increase in autonomy is supposed to lead to an increase in productiveness or justice—the former, because outside control may be less efficient, and the latter, because a dominant influence on the scientific process may constitute participatory injustice. However, as my discussion of the examples show, the relationship of scientific autonomy on one side and productiveness and justice on the other is ambivalent. Consider the example of ignored genes, or certain forms of publication bias: these are cases where the internal logic of scientific institutions—for example, what projects are the least risky for a successful scientific career—may lead to problematic imbalance, possibly in terms of both the main criteria. Still, this also shows why naive conceptions of scientific autonomy are problematic: are the necessities of scientific careers internal because they reflect the needs and ambitions of scientists, or are they external because they are the result of pressure on the individuals?

This also connects directly to the concept of conflicts of interest: Scientists and scientific institutions may have conflicting interests which, in the end, shape the distribution of research. In subsection 3.2.2, I have argued that the discussion about conflicts of interest in the meta-scientific discourse is too vague to be directly applied to cases of imbalance. I argued that it is especially problematic if we consider conflicts of interest as arising between a primary interest and secondary inadmissible ones because this difference is rarely explained in detail, especially in the context of scientific research. The framework presented in the last chapters contributes to the understanding of conflicts of interest in multiple ways. Firstly, we can be more explicit about what is problematic about conflicts of interest: violations of both criteria of imbalance and their respective sub-criteria can be applied to cases where we suspect conflicts of interest. They may either lead to erroneous beliefs, ignorance of important aspects of the subject matter, bad practical decisions, or they may in fact be unjust, where the prevailing interest of one party overrides the legitimate interests of other parties. Secondly, in my discussion of productiveness, my concept of relevance showed how actual, hypothetical and instrumental interests as well as curiosity and practical problems interact. Thirdly, the discussion of distributive justice revealed a possible explanation for the distinction of primary and secondary interests: as I have argued there, basic goods, related to human rights such as a right to health, and public goods, related to the functioning of public institutions, trump private interests, which are only beneficial to the non-basic interests of individuals.

As for the concept of bias, I have already argued in subsection 3.2.2 that imbalance as a problem with the SoRs can be both a result of bias and a cause of bias. My discussion on SoRs and imbalance has revealed an important aspect of the latter direction: Relying on Daniel Steel (2018), I argued that bias in science is not only about biased results, but also about the biasing effects the results may have further down the line, even if no individual finding can be said to be biased. Different from Steel, I have argued that this is not simply a matter of communication, where individual findings may be phrased in a way that may lead to misunderstandings. Instead, when we consider the overall SoR, especially lacunae can bias subsequent inferences even if the recipients of the research fully understand the results.

The implications of the concept of imbalance for objectivity are perhaps the most wide-ranging. While it does not seem strange, for any of the examples, that something about the examples that constitute imbalance is problematic also in terms of objectivity, the criteria of productiveness and justice are not usually considered elements of scientific objectivity.

As far as productiveness goes, the connection between alethic risk and objectivity may not come as a surprise; it will, however, be more controversial to include practical risk and efficiency as criteria that are related to objectivity. The plausibility of the former relation hinges to a considerable degree on the question if practical decisions can be objective: if they can, and if we can interpret the objectivity of a SoR in terms of how useful it is for making such objective decisions, then practical considerations do matter for objectivity. Efficiency, in contrast, may be one aspect where imbalance and objectivity diverge: If at all, a lack of efficiency may cast doubt on the question if the allocation of research resources was conducted in an objective manner. This, however, would indicate a procedural concept of objectivity. But a concept of objectivity that is a counter term to imbalance would have to be product-oriented, as imbalance itself is about a specific product—the SoR.

Even more interesting is the implication that justice and fairness may be elements of objectivity: In influential philosophical overviews of different theories about the concept of scientific objectivity such as Douglas (2004) or Reiss and Sprenger (2017), you will find but one very tangential reference to fairness or epistemic justice.[11][12] The kind of injustice that constitutes imbalance and can, therefore, be linked to a form of objectivity, is distributive justice. This line of thought stems from the special vantage point of this dissertation, which focused on distributions of research, and not on individual results. In contrast, many existing accounts of objectivity focus on individual scientific findings and the processes which may justify a claim to objectivity. It is therefore not surprising that these accounts do not focus on or even delineate objectivity from justice, as that might first appear to be some kind of category error. Even for individual findings, however, fairness can be an

11. Also this one reference is not about justice a requirement or element of objectivity, but as something opposed to it: "In most views, the objectivity and authority of science is not threatened by epistemic, but only by contextual (non-cognitive) values. Contextual values are moral, personal, social, political and cultural values such as pleasure, **justice and equality** [my emphasis], conservation of the natural environment and diversity." (Reiss and Sprenger 2017).

12. I cannot claim to be the very first philosopher to have suggested a connection between objectivity and fairness or justice. Naomi Scheman (2015), for one, writes that feminist philosophers of science and connected projects should argue for "the dependence of objectivity on the conditions of social justice" (ibid., 229). For her, however, this is connected to the larger claim that objectivity should depend on trustworthiness as a whole, which is a position I explicitly reject: objectivity is a precondition of trustworthiness, not the other way around.

important quality. It does not seem inappropriate to ask: Was the acceptance of this result based on a fair distribution of research, respecting the needs and viewpoints of the people affected by this claim to knowledge? And even further: For this single result, who benefits from this research, who does not, and is this distribution of benefits just? In this context, yet another meaning of independence comes in, linking objectivity and fairness. Consider some of the uses of "independence" in the context of information in expressions like "independent press" or "independent investigation". While, as I have argued before, independence guarantees neither objectivity nor fairness, it would not be surprising to see the qualifier "independent" replaced with either of the other concepts, indicating our expectations: an independent investigation, for example, could be expected to be both an objective investigation, or a fair investigation—fair, that is, to those affected by the outcome, not putting anyone at an unfair advantage or disadvantage. This additional connection between imbalance and independence also shows that in contexts other than science, objectivity and fairness are sometimes tightly linked. Consider, for example, the field of journalism. In Schudson (2001)—a treatise of objectivity norms connected to journalism in the USA—fairness and objectivity are used interchangeably. In Wien (2005)—an overview of different definitions of objectivity in journalism—this is made even more explicit. Regarding one such account of objectivity, Wien writes:

> As can be seen in the quotation, objectivity and fairness are synonyms. The most inter-esting thing here is that fairness can be graduated: one can be more or less fair. And that since there is an equivalence between being fair and being objective, one can thus also be more or less objective. (ibid., 9)

There are some interesting parallels between my concept of imbalance and the use of these concepts in journalism: not only do they connect fairness and objectivity, but like in my account, these concepts are presented as gradual, rather than binary criteria. Other than these authors, however, I do not equate objectivity and fairness: On one side, objectivity and imbalance are broader than justice, also including alethic if not practical risk and efficiency. But on the other side, there is also more to justice than what is covered by imbalance: Objectivity, even when expressed in procedural terms, is focused on some product that can be thought to be reliable, can be made use of as an epistemic resource. There are, of course, philosophers who support a procedural account of objectivity. With Douglas (2004) and also the "transformative criticism"-account of objectivity (cf. Longino 1990) we have encountered two such positions. However, I agree with Goldman (2003), who argues that, in the end, also authors like Longino implicitly judge the procedures in terms of their effects on the outcomes:

> Longino proceeds to emphasize how objectivity requires the avoidance of partiality or subjective preference, and argues that this avoidance is best secured by criticisms from the scientific community, which is precisely what her four criteria are intended to promote. It seems clear, then, that her general methodological framework is a consequentialist one: the virtue of various social practices is their tendency to promote greater impartiality and nonarbitrariness. (ibid., 78)

This, however, is different for injustice: When we talk about fairness as a procedural concept—such as in the case of participatory injustice—, the effects on the distribution of epistemic goods, and thus on the objectivity of a resource such as a finding or the overall SoR, is only instrumental; the intrinsic—and thus for some, the primary—wrong of participatory injustice is about an insult to individuals or groups in their capacities as knowers. This aspect of justice is incompatible with the usage of objectivity. As the previous discussion in this chapter has shown, intrinsic participatory injustice is also no criterion for imbalance, but rather, provides a reason to criticize the scientific process.[13]

Seeing that I have mentioned the concept of independence in connection with objectivity, one might object that rather than justice or fairness, avoiding certain influences on the SoR should be the way to explicate objectivity. But as critics of a value-free ideal have argued, avoiding all subjective influence altogether is largely either undesirable or impossible. Replacing this kind of objectivity based on value-freedom or independence with one rooted in objectivity avoids this pitfall, as justice

13. Interestingly enough, this difference also shows one point where trustworthiness is decisively distinct from objectivity: certainly, intrinsic procedural injustice can be a source of distrust, but, as just argued, not in itself a reason for denying objectivity.

admits the influence of particular perspectives, as long as they don't skew the outcomes towards an unfair disadvantage for any particular group. Still, some may worry that a concept of objectivity that includes distributive justice as one of its constituent elements may be too normatively laden to account for some of the meanings of objectivity. Especially when we use phrases such as "Objectivity as Faithfulness to Facts" (Reiss and Sprenger 2017) or "getting at objects in the world" (Douglas 2004, 472), this thought has its merits. When we think of the examples presented in the historical analysis of objectivity in Daston and Galison (2010), applying a concept of distributive justice seems far-fetched: Can we, instead of talking about objectivity, talk about *fair* drawings of liquid drop experiments (cf. ibid., 11-16) or specimens of botany or anatomy, and still refer to the same quality? However, in other contexts, my proposed idea of objectivity should appear less alien. While defenders of a value-free ideal would perhaps want to avoid fairness as a criterion belonging to objectivity when talking about the justification of individual claims, even they would surely admit that fairness has an important role to play when determining the research agenda—which is pivotal for the SoR. The consequence of this line of thought, therefore, should not be to abandon a view of objectivity as productiveness and fairness. Instead, we should recognize that although objectivity may always be related to some core ideas—perhaps to the usefulness of an epistemic resource—this core would only be the lowest common denominator, transmitted through a line of historically grown and changing concepts of objectivity, which—as Douglas (2004) argues—are irreducible to each other. The concept of objectivity that is linked to imbalance, and thus to justice, is only one possible use, but, as I hope to have shown, it is a useful one. It is the concept of objectivity we seek when we want to know, if, in our complex world, with science informing policy and individual judgment on far-reaching issues, relying on the SoR will lead to objective opinions and objective decisions.

7.2.3 Outlook

The discussion in this dissertation had a broad scope, addressing both epistemological and ethical issues: anchored in an interest in imbalances in the SoR, I have explored topics from alethic risk to epistemic injustice. I have covered a lot of ground over the course of this analysis, and I hope I have provided a convincing argument for showing an interest in problems with the distribution of research. However, there is still a lot of work to be done: Clearly, the examples of imbalance themselves warrant further investigation. Firstly, the list of examples provided here is open-ended, and it would be very interesting to compare how the framework can deal with new cases, for example from other areas less close to health and environment, or even practical applications altogether. Also, my analysis has been focused on examples from the sciences. The SoR, however, is a concept that could also be applied in other fields, such as the humanities, or philosophy itself. Also for the cases already considered, the application of the criteria in this chapter has shown that the explanations suggested by the authors of the original examples are not the only possible ones; for each of them, we can explore if these other lines of criticism can be substantiated as well. This, then, is not only about philosophical work, but calls for investigations by many other meta-scientific disciplines, be it sociology of science, psychology, economics, law, or political science. The framework provided in this dissertation does not directly suggest easy to use measures of the individual criteria that could be applied in empirical investigations of imbalance. However, what it does do is broaden the scope of possible consequences of imbalance that can and should be considered.

But not only the examples of imbalance themselves merit further investigation. In all philosophical fields that I have touched upon in the individual chapters, there remain some open theoretical questions: Given my reservations against simple, easily quantifiable measures of alethic risk, how do we answer the challenge of comparing SoR concerning the promotion of truth? How would we determine an overall atlas of scientific significance that could integrate the perspectives from different SoRs and the internal relevance of the research questions contained within? What theory of justice is most convincing when applied to research as a good that can inform our decisions? These questions and many others could deepen our understanding of defects with scientific research on the level of the SoR. In the last section, I have already indicated some lines of thought concerning the wider theoretical background: Especially for the concept of objectivity, I propose we should focus even more on the practical consequences of research as an objective source of information. In this context, I believe, we philosophers of science would do well to look beyond our own subject, and question how familiar concepts like objectivity have been framed in other discourses, for example, about journalism and the media in general. Generally, I believe philosophers of science should show more interest in the dissemination of research: in the publishing process, in how the SoR gets reconstructed, and how this information is taken up within and beyond the inner sphere of scientific research.

Bibliography

AIFA. 2010. “Feasibility and Challenges of Independent Research on Drugs: The Italian Medicines Agency (AIFA) Experience.” *European Journal of Clinical Investigation* 40 (1): 69–86.

Anderegg, William R. L. 2010. “Moving Beyond Scientific Agreement: An Editorial Comment on “Climate Change: A Profile of Us Climate Scientists’ Perspectives”.” *Climatic Change* 101 (3): 331–337.

Anderson, Elizabeth. 2012. “Epistemic Justice as a Virtue of Social Institutions.” *Social Epistemology* 26 (2): 163–173.

Ankeny, Rachel A. 2014. “The Overlooked Role of Cases in Casual Attribution in Medicine.” *Philosophy of Science* 81 (5): 999–1011.

Autzen, Bengt. 2016. “Significance Testing, P-Values and the Principle of Total Evidence.” *European Journal for Philosophy of Science* 6 (2): 281–295.

Baier, Annette. 1986. “Trust and Antitrust.” *Ethics* 96 (2): 231–260.

Becker, Lawrence C. 1996. “Trust as Noncognitive Security about Motives.” *Ethics* 107 (1): 43–61.

Betz, Gregor. 2013. “In Defence of the Value Free Ideal.” *European Journal for Philosophy of Science* 3 (2): 207–220.

Bias | Origin and Meaning of Bias. n.d. In *Online Etymology Dictionary,* by Douglas Harper. Accessed September 26, 2019. https://www.etymonline.com/word/bias.

Biddle, Justin B., and Rebecca Kukla. 2017. "The Geography of Epistemic Risk." In *Exploring Inductive Risk,* edited by Kevin C. Elliott and Ted Richards, 215–237. Oxford University Press.

Biddle, Justin B., and Anna Leuschner. 2015. "Climate Skepticism and the Manufacture of Doubt: Can Dissent in Science Be Epistemically Detrimental?" *European Journal for Philosophy of Science* 5 (3): 261–278.

Brown, James Robert. 2001. *Who Rules in Science? An Opinionated Guide to the Wars.* Cambridge, Mass.: Harvard University Press.

———. 2008. "The Community of Science®." In *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited,* edited by Martin Carrier, Don Howard, and Janet Kourany, 189–216. University of Pittsburgh Press.

———. 2016. "Patents and Progress." *Perspectives on Science* 24 (5): 505–528.

———. 2017. "Socializing Medical Research." In *Current Controversies in Values and Science,* 1st ed., edited by Kevin C. Elliott and Daniel Steel, 147–160. Routledge.

Bueter, Anke. 2015. "The Irreducibility of Value-Freedom to Theory Assessment." *Studies in History and Philosophy of Science Part A* 49:18–26.

Carel, Havi, and Ian James Kidd. 2017. "Epistemic Injustic in Medicine and Healthcare." In *The Routledge Handbook of Epistemic Injustice,* edited by Ian James Kidd, José Medina, and Gaile Pohlhaus, 336–346. Routledge Handbooks in Philosophy. London New York: Routledge, Taylor & Francis Group.

Carrier, Martin. 2008. "Science in the Grip of the Economy: On the Epistemic Impact of the Commercialization of Research." In *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited,* edited by Martin Carrier, Don Howard, and Janet Kourany, 217–234. University of Pittsburgh Press.

Cartwright, Nancy. 2012. "Presidential Address: Will This Policy Work for You? Predicting Effectiveness Better: How Philosophy Helps." *Philosophy of Science* 79 (5): 973–989.

Chalmers, Iain, Larry V. Hedges, and Harris Cooper. 2002. "A Brief History of Research Synthesis." *Evaluation & the Health Professions* 25 (1): 12–37.

Chan, An-Wen, Fujian Song, Andrew Vickers, Tom Jefferson, Kay Dickersin, Peter C Gøtzsche, Harlan M Krumholz, Davina Ghersi, and H Bart van der Worp. 2014. "Increasing Value and Reducing Waste: Addressing Inaccessible Research." *The Lancet* 383 (9913): 257–266.

Christian, Alexander. 2017. "On the Suppression of Medical Evidence." *Journal for General Philosophy of Science* 48 (3): 395–418.

Coady, David. 2010. "Two Concepts of Epistemic Injustice." *Episteme* 7 (2): 101–113.

———. 2017. "Epistemic Injustice as Distributive Injustice." In *The Routledge Handbook of Epistemic Injustice,* edited by Ian James Kidd, José Medina, and Gaile Pohlhaus, 61–68. Routledge Handbooks in Philosophy. London New York: Routledge, Taylor & Francis Group.

Cook, John, Naomi Oreskes, Peter T. Doran, William R. L. Anderegg, Bart Verheggen, Ed W. Maibach, J. Stuart Carlton, et al. 2016. "Consensus on Consensus: A Synthesis of Consensus Estimates on Human-Caused Global Warming." *Environmental Research Letters* 11, no. 4 (1, 2016): 048002.

Core Writing Team, R. K. Pachauri, and L. A. Meyer, eds. 2015. *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Geneva, Switzerland: Intergovernmental Panel on Climate Change.

Daston, Lorraine, and Peter Galison. 2010. *Objectivity.* Paperback. New York, NY: Zone Books.

De Winter, Jan, and Laszlo Kosolosky. 2014. "Health, Food, and Science." *Logique et Analyse* 57 (228): 701–726.

Douglas, Heather. 2000. "Inductive Risk and Values in Science." *Philosophy of Science* 67 (4): 559–579.

Douglas, Heather. 2004. "The Irreducible Complexity of Objectivity." *Synthese* 138 (3): 453–473.

———. 2012. "Weighing Complex Evidence in a Democratic Society." *Kennedy Institute of Ethics Journal* 22 (2): 139–162.

Dworkin, Ronald. 2002. *Sovereign Virtue: The Theory and Practice of Equality.* 4. print. Cambridge, Mass.: Harvard Univ. Press.

Edwards, Aled M., Ruth Isserlin, Gary D. Bader, Stephen V. Frye, Timothy M. Willson, and Frank H. Yu. 2011. "Too Many Roads Not Taken." *Nature* 470 (7333): 163–165.

Elliott, Kevin C., and Daniel J. McKaughan. 2009. "How Values in Scientific Discovery and Pursuit Alter Theory Appraisal." *Philosophy of Science* 76 (5): 598–611.

European Science Foundation. 2011. *The European Code of Conduct for Research Integrity.*

Field, Christopher B., Vicente R. Barros, David Jon Dokken, Katharine J. Mach, and Michael D. Mastrandrea, eds. 2014. "Preface." In *Climate Change 2014 Impacts, Adaptation, and Vulnerability,* ix–x. Cambridge: Cambridge University Press.

Flory, James H., and Philip Kitcher. 2004. "Global Health and the Scientific Research Agenda." *Philosophy & Public Affairs* 32 (1): 36–65.

Frickel, Scott, Sahra Gibbon, Jeff Howard, Joanna Kempner, Gwen Ottinger, and David J. Hess. 2010. "Undone Science: Charting Social Movement and Civil Society Challenges to Research Agenda Setting." *Science, Technology, & Human Values* 35 (4): 444–473.

Fricker, Miranda. 2007. *Epistemic Injustice.* Oxford University Press.

———. 2013. "Epistemic Justice as a Condition of Political Freedom?" *Synthese* 190 (7): 1317–1332.

———. 2017. “Evolving Concepts of Epistemic Injustice.” In *The Routledge Handbook of Epistemic Injustice,* edited by Ian James Kidd, José Medina, and Gaile Pohlhaus, 53–60. Routledge handbooks in philosophy. London New York: Routledge, Taylor & Francis Group.

Goldman, Alvin I. 2003. *Knowledge in a social world.* Reprint. Oxford: Clarendon Press.

Grant, Maria J., and Andrew Booth. 2009. “A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies.” *Health Information & Libraries Journal* 26 (2): 91–108.

Grasswick, Heidi. 2017. “Epistemic Injustice in Science.” In *The Routledge Handbook of Epistemic Injustice,* edited by Ian James Kidd, José Medina, and Gaile Pohlhaus, 313–323. Routledge Handbooks in Philosophy. London New York: Routledge, Taylor & Francis Group.

Gruening, Thilo, Anna B. Gilmore, and Martin McKee. 2006. “Tobbaco Industry Influence on Science and Scientists in Germany.” *American Journal of Public Health* 96 (1): 20–32.

Hardin, Russell. 1991. “Trusting Persons, Trusting Institutions.” In *Strategy and Choice,* 2. print, edited by Richard J. Zeckhauser, 185–209. Cambridge, Mass.: MIT Press.

———. 1996. “Trustworthiness.” *Ethics* 107 (1): 26–42.

———. 1998. “Trust in Government.” In *Trust and Governance,* edited by Valerie A. Braithwaite and Margaret Levi. The Russell Sage Foundation series on trust 1. New York: Russell Sage Foundation.

Hardwig, John. 1991. “The Role of Trust in Knowledge.” *The Journal of Philosophy* 88 (12): 693–708.

Higgins, Julian PT, and Sally Green, eds. 2008. *Cochrane Handbook for Systematic Reviews of Interventions.* Chichester, UK: John Wiley & Sons, Ltd.

Huber, Franz. 2008. “Belief and Degrees of Belief.” In *Degrees of Belief,* edited by Franz Huber and Christoph Schmidt-Petri, 1–33. Springer Science & Business Media, December 21, 2008.

Irzik, Gürol, and Faik Kurtulmus. 2018. “What is Epistemic Public Trust in Science?” *The British Journal for the Philosophy of Science* (22, 2018): 1–22.

John, S. D. 2014. “Risk, Contractualism, and Rose’s” Prevention Paradox”.” *Social Theory and Practice,* 28–50.

Jordan, Catherine, Susan Gust, and Naomi Scheman. 2011. “The Trustworthiness of Research: The Paradigm of Community-Based Research.” In *Shifting Ground: Knowledge and Reality, Transgression and Trustworthiness,* by Naomi Scheman, 170–190. Oxford University Press.

Kearns, Cristin E., Stanton A. Glantz, and Laura A. Schmidt. 2015. “Sugar Industry Influence on the Scientific Agenda of the National Institute of Dental Research’s 1971 National Caries Program: A Historical Analysis of Internal Documents.” Edited by Simon Capewell. *PLOS Medicine* 12, no. 3 (10, 2015): e1001798.

Kelly, Thomas. 2016. “Evidence.” In *The Stanford Encyclopedia of Philosophy,* Winter 2016, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Kidd, Ian James, and Havi Carel. 2017. “Epistemic Injustice and Illness.” *Journal of Applied Philosophy* 34 (2): 172–190.

Kitcher, Philip. 1993. *The Advancement of Science: Science Without Legend, Objectivity Without Illusions.* 1. Pb. publ. New York: Oxford Univ. Press.

———. 2001. *Science, Truth, and Democracy.* Oxford studies in philosophy of science. New York: Oxford Univ. Press.

———. 2011. *Science in a Democratic Society.* Amherst, NY: Prometheus Books.

———. 2015. “Pragmatism and Progress.” *Transactions of the Charles S. Peirce Society* 51 (4): 475–494.

Kuhn, Thomas S. 1977. “Objectivity, Value Judgment, and Theory Choice.” In *The Essential Tension,* 320–339. University of Chicago Press.

Kurtulmus, Faik, and Gürol Irzik. 2017. “Justice in the Distribution of Knowledge.” *Episteme* 14 (2): 129–146.

Lakatos, Imre. 1976a. “Falsification and the Methodology of Scientific Research Programmes.” In *Can Theories be Refuted?,* edited by Sandra G. Harding, 8–101. Dordrecht: Springer Netherlands.

———. 1976b. “History of Science and Its Rational Reconstructions.” In *Can Theories be Refuted?,* edited by Sandra G. Harding, 102–138. Dordrecht: Springer Netherlands.

Leiserowitz, Anthony A., Edward W. Maibach, Connie Roser-Renouf, Nicholas Smith, and Erica Dawson. 2013. “Climategate, Public Opinion, and the Loss of Trust.” *American Behavioral Scientist* 57 (6): 818–837.

Leslie, S.-J., A. Cimpian, M. Meyer, and E. Freeland. 2015. “Expectations of Brilliance Underlie Gender Distributions Across Academic Disciplines.” *Science* 347, no. 6219 (16, 2015): 262–265.

Leuschner, Anna. 2018. “Is It Appropriate to ‘Target’ Inappropriate Dissent? On the Normative Consequences of Climate Skepticism.” *Synthese* 195 (3): 1255–1271.

Lewandowsky, Stephan, Naomi Oreskes, James S. Risbey, Ben R. Newell, and Michael Smithson. 2015. “Seepage: Climate Change Denial and Its Effect on the Scientific Community.” *Global Environmental Change* 33:1–13.

Lewandowsky, Stephan, James S. Risbey, and Naomi Oreskes. 2016. “The “Pause” in Global Warming: Turning a Routine Fluctuation into a Problem for Science.” *Bulletin of the American Meteorological Society* 97 (5): 723–733.

Lo, Bernard, and Marilyn Jane Field, eds. 2009. *Conflict of Interest in Medical Research, Education, and Practice.* Washington, DC: National Acad. Press.

Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry.* Princeton, NJ: Princeton Univ. Pr.

———. 2008. "Values, Heuristics, and the Politics of Knowledge." In *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited,* edited by Martin Carrier, Don Howard, and Janet A. Kourany, 68–86. Pittsburgh, Pa: University of Pittsburgh Press.

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J. F. Gonzalez Rouco, E. Jansen, et al. 2013. "Information from Paleoclimate Archives." In *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change,* edited by T.F. Stocker, D. Qin, Gian-Kasper Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, 383–464. Cambridge Univ. Press, Cambridge, UK, / New York.

Mayer, Roger C., James H. Davis, and F. David Schoorman. 1995. "An Integrative Model of Organizational Trust." *The Academy of Management Review* 20 (3): 709–734.

Melo-Martín, Inmaculada de, and Kristen Intemann. 2018. *The Fight Against Doubt: How to Bridge the Gap Between Scientists and the Public.* New York, NY: Oxford University Press.

Mill, John Stuart. 2015. *On Liberty, Utilitarianism and Other Essays.* New edition. Edited by Mark Philip and Frederick Rosen. Oxford: Oxford University Press.

Miller, David. 2017. "Justice." In *The Stanford Encyclopedia of Philosophy,* Fall 2017, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Nair, P. K. Ramachandran, and Vimala D. Nair. 2014. "Organization of a Research Paper: The IMRAD Format." In *Scientific Writing and Communication in Agriculture and Natural Resources,* 13–25. Cham: Springer International Publishing.

Oreskes, N. 2004. "Beyond the Ivory Tower: The Scientific Consensus on Climate Change." *Science* 306, no. 5702 (3, 2004): 1686–1686.

Oreskes, Naomi, and Erik M. Conway. 2012. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming.* Paperback. London: Bloomsbury.

Petit, J. R., J. Jouzel, D. Raynaud, N. I. Barkov, J.-M. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davisk, and G. Delaygue. 1999. "Climate and Atmospheric History of the Past 420,000 Years from the Vostok Ice Core, Antarctica."

Pielke, Roger A., Jr. 2007. *The Honest Broker: Making Sense of Science in Policy and Politics.* 1st ed. Cambridge University Press.

Pigman, W., and E. B. Carmichael. 1950. "An Ethical Code for Scientists." *Science* 111, no. 2894 (16, 1950): 643–647.

Public, Adj. and N. 2007. In *OED Online.* Oxford University Press. Accessed August 15, 2019. https://www.oed.com/view/Entry/154052.

Railton, Peter. 1981. "Probability, Explanation, and Information." *Synthese* 48 (2): 233–256.

Rawls, John. 1971. *A Theory of Justice.* Orig. ed., reprint. Cambridge, Mass.: Belknap Press.

Reiss, Julian. 2017. "Meanwhile, Why Not Biomedical Capitalism?" In *Current Controversies in Values and Science,* 1st ed., edited by Kevin C. Elliott and Daniel Steel, 161–175. Routledge.

Reiss, Julian, and Philip Kitcher. 2009. "Biomedical Research, Neglected Diseases, and Well-Ordered Science." *THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia* 24:263–282.

Reiss, Julian, and Jan Sprenger. 2017. "Scientific Objectivity." In *The Stanford Encyclopedia of Philosophy,* Winter 2017, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Resnik, David B. 2007. *The Price of Truth: How Money Affects the Norms of Science.* Practical and professional ethics series. New York: Oxford Univ. Press.

———. 2011. "Scientific Research and the Public Trust." *Science and Engineering Ethics* 17 (3): 399–409.

Robock, Alan. 2008. "20 Reasons Why Geoengineering May Be a Bad Idea." *Bulletin of the Atomic Scientists* 64, no. 2 (1, 2008): 14–18.

Rolin, Kristina. 2015. "Values in Science: The Case of Scientific Collaboration." *Philosophy of Science* 82 (2): 157–177.

Rooney, Phyllis. 1992. "On Values in Science: Is the Epistemic/Non-Epistemic Distinction Useful?" In *PSA: Proceedings of the biennial meeting of the philosophy of science association,* 13–22.

Rosenberg, Stacy, Arnold Vedlitz, Deborah F. Cowman, and Sammy Zahran. 2010. "Climate Change: A Profile of Us Climate Scientists' Perspectives." *Climatic Change* 101 (3): 311–329.

Rosenthal, Robert. 1979. "The File Drawer Problem and Tolerance for Null Results." *Psychological Bulletin* 86 (3): 638–641.

Rudner, Richard. 1953. "The Scientist Qua Scientist Makes Value Judgments." *Philosophy of science* 20 (1): 1–6.

Sarewitz, Daniel. 2016. "Saving Science." *The New Atlantis,* no. 49, 4–40.

Saul, Jennifer. 2013. "Implicit Bias, Stereotype Threat, and Women in Philosophy." In *Women in Philosophy: What Needs to Change,* edited by Hutchison Katrina and Fiona Jenkins, 39–60.

Schekman, Randy. 2013. "How journals like Nature, Cell and Science are damaging science | Randy Schekman." *The Guardian* (9, 2013). Accessed October 11, 2019. https://www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science.

Scheman, Naomi. 2015. "Epistemology Resuscitated: Objectivity as Trustworthiness." In *Shifting Ground: Knowledge and Reality, Transgression and Trustworthiness,* 207–231. Oxford University Press.

Schudson, Michael. 2001. "The Objectivity Norm in American Journalism." *Journalism* 2 (2): 149–170.

Sismondo, Sergio. 2008. "How Pharmaceutical Industry Funding Affects Trial Outcomes: Causal Structures and Responses." *Social Science & Medicine* 66 (9): 1909–1914.

Sollaci, Luciana B., and Mauricio G. Pereira. 2004. "The Introduction, Methods, Results, and Discussion (imrad) Structure: A Fifty-Year Survey." *Journal of the Medical Library Association: JMLA* 92 (3): 364–367.

Song, F., A. J. Eastwood, S. Gilbody, L. Duley, and A. J. Sutton. 2000. "Publication and Related Biases." *Health Technology Assessment* 4 (10).

Song, F., S. Parekh, L. Hooper, Yk. Loke, J. Ryder, Aj. Sutton, C. Hing, Cs. Kwok, C. Pang, and I. Harvey. 2010. "Dissemination and Publication of Research Findings: An Updated Review of Related Biases." *Health Technology Assessment* 14 (8).

Steel, Daniel. 2018. "If the Facts Were Not Untruths, Their Implications Were: Sponsorship Bias and Misleading Communication." *Kennedy Institute of Ethics Journal* 28 (2): 119–144.

Stegenga, Jacob. 2011. "Is Meta-Analysis the Platinum Standard of Evidence?" *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 42 (4): 497–507.

———. 2018. *Medical Nihilism.* Oxford University Press.

Sterling, Theodore D. 1959. "Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance–or Vice Versa." *Journal of the American Statistical Association* 54 (285): 30–34.

Todorović, Ljubomir. 2003. "Original (Scientific) Paper: The IMRaD Layout." *Archive of Oncology* 11 (3): 203–205.

Van Norman, Gail A. 2016. "Drugs, Devices, and the FDA: Part 1." *JACC: Basic to Translational Science* 1 (3): 170–179.

Whitbeck, Caroline. 1995. "Trustworthy Research—an Editorial Introduction." *Science and Engineering Ethics* 1 (4): 322–328.

Whyte, Kyle Powys, and Robert P. Crease. 2010. "Trust, Expertise, and the Philosophy of Science." *Synthese* 177 (3): 411–425.

Wien, Charlotte. 2005. "Defining Objectivity within Journalism: An Overview." *Nordicom Review* 26 (2): 3–15.

Wilholt, Torsten. 2010. "Scientific Freedom: Its Grounds and Their Limitations." *Studies in History and Philosophy of Science Part A* 41 (2): 174–181.

———. 2012. *Die Freiheit der Forschung: Begründungen und Begrenzungen.* Orig.-Ausgabe, 1. Auflage. Suhrkamp-Taschenbuch Wissenschaft 2040. Berlin: Suhrkamp.

———. 2013. "Epistemic Trust in Science." *The British Journal for the Philosophy of Science* 64, no. 2 (1, 2013): 233–253.

———. 2016. "Collaborative Research, Scientific Communities, and the Social Diffusion of Trustworthiness." In *The Epistemic Life of Groups,* edited by Michael S. Brady and Miranda Fricker, 218–234. Oxford University Press.