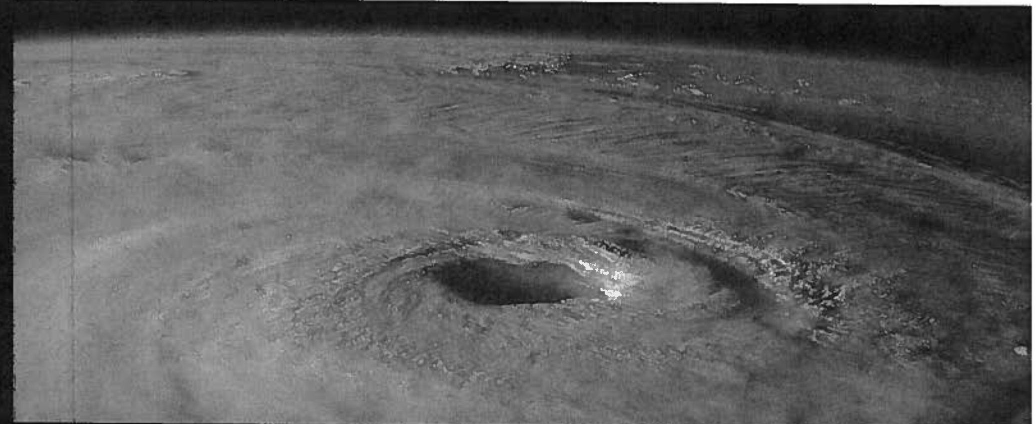
new waves
in philosophy



new waves in philosophy of science

edited by

p.d. magnus & jacob busch

New Waves in Philosophy
Series Standing Order ISBN 978–0–230–53797–2 (hardcover)
Series Standing Order ISBN 978–0–230–53798–9 (paperback)
(*outside North America only*)

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in case of difficulty, write to us at the address below with your name and address, the title of the series and the ISBN quoted above.

Customer Services Department, Macmillan Distribution Ltd, Houndmills, Basingstoke, Hampshire RG21 6XS, England

New Waves in Philosophy of Science

Edited by

P. D. Magnus
University of Albany, USA

and

Jacob Busch
University of St Andrews, UK

palgrave
macmillan

First published 2010 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries

ISBN 978–0–230–22263–2 hardback
ISBN 978–0–230–22264–9 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
19 18 17 16 15 14 13 12 11 10

Printed and bound in Great Britain by
CPI Antony Rowe, Chippenham and Eastbourne

Contents

List of Figures

Series Editors' Preface

New Waves in Philosophy Series

The aim of this series is to gather the young and up-and-coming scholars in philosophy to give their view of the subject now and in the years to come, and to serve a documentary purpose, i.e. 'this is what they said then, and this is what happened'. It will also provide a snapshot of cutting-edge research that will be of vital interest to researchers and students working in all subject areas of philosophy.

The goal of the series is to have a New Waves volume in every one of the main areas of philosophy. We would like to thank Palgrave Macmillan for taking on this project in particular, and the entire *New Waves in Philosophy* series in general.

Vincent F. Hendricks and Duncan Pritchard
Editors

Notes on Contributors

Justin Biddle is Assistant Professor in the School of Public Policy, Georgia Institute of Technology, USA. He received his PhD in the History and Philosophy of Science from the University of Notre Dame in 2006. His primary research interests are in the philosophy of science, social epistemology, and the ethics of science. He has published in journals such as *Social Epistemology*.

Otávio Bueno is a Professor at the Department of Philosophy, University of Miami, USA, before which he was an Assistant Professor at California State University and at the University of South Carolina where he was also made Associate Professor. He received his PhD from the University of Leeds in 1999. He has done extensive work in the area of Philosophy of Mathematics as well as in Philosophy of Science and has published on these issues in journals such as *Synthese*, *Erkenntnis*, *Studies in the History and Philosophy of Science*, *Philosophy of Science* and many more. Amongst his many other achievements he is co-editor of the volume *New Waves in Philosophy of Mathematics*.

Jacob Busch (PhD University of Auckland, 2006) is a research fellow in philosophy at the University of St Andrews and associate research fellow with the NAMICONA Research Centre at the University of Aarhus. Before moving to St Andrews he was a visiting researcher with the HPS, at Leeds (2005–2006). His work has appeared in journals such as *International Studies in the Philosophy of Science* and *Studies in History and Philosophy of Science*. He specializes in the metaphysical aspect of the realist/anti-realist debate in the philosophy of science.

Anjan Chakravartty is Associate Professor, University of Toronto Institute for the History and Philosophy of Science and Technology, and Department of Philosophy, Canada. He received a BSc in Biophysics from the University of Toronto and a PhD from the University of Cambridge in 2001. From 2000–2002 he was a Research Fellow at King's College, University of Cambridge. He is the author of numerous articles published in journals such as *Philosophy of Science*, *Synthese*, *Studies in the History and Philosophy of Science*, and the book *A Metaphysics for Scientific Realism: Knowing the Unobservable* (2007).

Gregory Frost-Arnold is Assistant Professor at the University of Nevada, Las Vegas, USA. He received his PhD from the University of Pittsburgh in 2006. Dr Frost-Arnold has published in journals such as *Philosophy of Science*, *Biology and Philosophy*, and *The Journal of Philosophical Logic*.

Kristen Intemann is an Assistant Professor at Montana State University, USA. She received her PhD from the University of Washington in 2004. She specializes in the areas of values and science and feminist philosophy of science. Dr Intemann has published in journals such as *Philosophy of Science* and *Science and Education*.

P. D. Magnus (PhD University of California San Diego, 2003) is an Assistant Professor in Philosophy at the University at Albany, State University of New York. Before moving to Albany, he was a visiting assistant professor at Bowdoin College (2003–2004). His work in the philosophy of science has appeared in journals such as *Philosophy of Science*, *The British Journal for the Philosophy of Science*, *Synthese*, and *Social Studies of Science*. Much of his work has concerned the alleged underdetermination of theory.

Jay Odenbaugh is Assistant Professor at Lewis and Clark College, USA. Before that he was a lecturer at University of California, San Diego. He received his PhD from the University of Calgary in 2001. He specializes in Philosophy of Biology and Environmental Ethics. He has published in journals such as *Philosophy of Science*, *Biology and Philosophy* and *Environmental Values*. He is a contributor to the *Encyclopedia for the Philosophy of Science* edited by Jessica Pfeifer and Sahotra Sarkar (2006).

Laura Perini is Assistant Professor at the Department of Philosophy, Pomona College, USA. She has an MA in Biology from the University of California, Los Angeles. She received her PhD from the University of California, San Diego in 2002. She has been a fellow at the Center for Philosophy of Science, University of Pittsburgh and a visiting fellow at Dartmouth College Humanities Institute, and she previously taught at Virginia Tech. She has published in journals such as *Philosophy of Science* and *Philosophy and Biology* and contributed to the *The Encyclopedia of Philosophy of Science* edited by Jessica Pfeifer and Sahotra Sarkar (2006).

Jessica Pfeifer is Assistant Professor at the University of Maryland, USA. She completed her PhD in Philosophy/Science Studies at the University of California, San Diego in 1999. The primary focus of her research has been on the nature of modal notions, and in particular the natural necessity involved in lawful relations. More recently she has focused on questions in the Philosophy of Biology. Pfeifer is co-editor (along with Sahotra Sarkar) of the two-volume *The Encyclopedia of Philosophy of Science* (2006). She has also been a Visiting Fellow and is currently an Associate Fellow at the Pittsburgh Center for Philosophy of Science.

Sherrilyn Roush is Associate Professor at the University of Berkeley, USA. Before that she was an assistant professor at Rice University. She has a BSc in Mathematics and she received her PhD from Harvard University in 1999. Roush has published in a number of journals, such as *Journal of Philosophy*, *Synthese*, *Philosophy of Science* and *Studies in History and Philosophy of Modern Physics*. She is also the author of the book *Tracking Truth: Knowledge, Evidence and Science* (2005).

Juha Saatsi has an MSc degree in Theoretical Physics from the University of Jyväskylä, Finland and completed his PhD in Philosophy at the University of Leeds, UK in 2006. From 2005–2006 he was a postdoctoral fellow with the University of Manchester and from 2006–2007 a postdoctoral fellow with the University of Leeds. He is currently a teaching fellow at the University of Leeds. Dr Saatsi is the author of numerous articles in such journals as *Philosophy of Science*, *British Journal for the Philosophy of Science*, and *Studies in History and Philosophy of Science*. He is also co-editor of *The Structural Foundations of Quantum Gravity* (2006).

Daniel Steel is Associate Professor at Michigan State University, USA. He has a BA in Anthropology and received his PhD from the University of Pittsburgh in 2002. He held a fellowship at the Center for Philosophy of Science at the University of Pittsburgh 2005. He specializes in Philosophy of Social Science as well as general Philosophy of Science and Philosophy of Biology. He has published in such journals as *British Journal for the Philosophy of Science*, *Synthese*, *Philosophy of Science*, and *Philosophy of the Social Sciences*. He is author of the forthcoming book *Across the Boundaries: Extrapolation and Causality in the Biological and Social Sciences*.

Michael Weisberg is Assistant Professor at the University of Pennsylvania, USA. Since 2006 he has been an associate at the Center for Philosophy of Science, University of Pittsburgh. He has a BSc in Chemistry from University of California, San Diego and he received his PhD from Leland Stanford Junior University in 2003. He has published in journals such as *Philosophy of Science*, *Studies in the History and Philosophy of Science*, and *British Journal for the Philosophy of Science*.

Eric Winsberg is Associate Professor at the University of South Florida, USA. He received his PhD from Indiana University in 1999. He held a postdoctoral fellowship at Northwestern University from 1999–2001 and has been a fellow at the Center for Interdisciplinary Studies, University of Bielefeld from 2006–2007. Since 2008 he has been a fellow at the Institute of Advanced Studies, University of Durham. He has published in journals such as *Synthese*, *Philosophy of Science*, and *Journal of Philosophy*. He is currently writing a new book: *Science in the Age of Computer Simulation*.

Introduction

P. D. Magnus and Jacob Busch

The explicit aim of volumes in this series is to collect contributions from young researchers likely to dominate the discipline; for this volume, the discipline in question is philosophy of science. It has been our privilege to edit such an audacious project, but it has also been a great challenge. We can only make educated guesses about the future. We cannot say with certainty which recent topics will be central to the discipline. Even selecting among areas which will probably remain central, there are competing desiderata: for example, to balance perennial topics against others which have more recently attracted the attention of philosophers.

A number of subjects have recurred in debates among philosophers of science: induction and the justificatory status of ampliative modes of inference; the role of explanation and its metaphysical status; causation; the problem of demarcation; the role of metaphysics in science; probability theory and its influence on formal approaches to confirmation; and scientific realism against various kinds of anti-realism. One might also add issues arising out of twentieth-century physics, such as the interpretation of quantum mechanics and the nature of spacetime theories. And it is no longer the case that a philosopher can claim to account for science when only really accounting for physics. So one might add issues that arise in biology, chemistry, climatology, psychology, sociology, and other sciences.

Philosophy of physics, philosophy of biology, and others have become autonomous specialties in their own right. There are conferences and journals dedicated to each, and we could without difficulty have collected papers for a volume exploring the 'new waves' in just one of them. So we are the first to admit that any volume covering philosophy of science *tout court* will leave out a great deal which might reasonably have been included.

We did not pick a roster of topics and invite each author to report on the state of the art in one area. That might have resulted in a stultifying collection of field reports. Instead, we invited contributors who are doing exciting work in diverse areas of philosophy of science. We explained the aim of the volume. And then we let them tell us what needed to be written, which debates are the important ones, and how those debates might best be resolved.

A consequence of our approach is that the table of contents is not an exhaustive roster of topics. Also, unfortunately, not all of the invited authors were ultimately able to contribute. So topics like causation, the role of experiment, and formal confirmation are regrettably absent. The field is large enough, however, that some omissions were inevitable. We apologize for them, all the same.

The alternative to our approach would have meant filling out the contents of the volume as if it were a shopping list, with topics that could be checked off as they were put in the basket. That would make sense if philosophy of science were nothing more than a mismatched agglomeration of all of these specialties. On that agglomeration conception, the authors in this volume are working in disparate fields. Yet, once asked to identify a pressing issue – something they are working on now that will still be worth thinking about a decade from now – none selected boutique topics. All of the topics addressed in this volume are within the purview of general philosophy of science, rather than being of interest only to sub-specialists within a fragment of a specialty. The perennial issues of realism and representation are presented here in their twenty-first-century incarnations. Where specific examples are discussed, they are used to illuminate more general questions. For example: electron micrographs raise questions about the nature of visual representation. Environmental science and climate change are test cases for accounts of the relation between values and science. A central issue in the philosophy of social science ultimately turns on our conception of science itself. And so on.

For the purpose of our introduction, we'll treat the topics in five broad groups: (1) scientific realism, (2) science and metaphysics, (3) representation, (4) the relation between science and values, and (5) science and social structure.

1. Realism

Scientific realism is a stock issue in the philosophy of science, motivated in its current form by the failure of logical positivism in the middle of

the last century. The subject has both benefited from and to some extent suffered from the amount of literature on the various aspects of realism. Today there is little agreement on what constitutes the most significant aspect of the debate, but epistemic concerns are inescapable; realists must rely on ampliative modes of inference.

Juha Saatsi maps out the territory, proposing that various ampliative strategies are best organized by the conceptual distinction between content-driven and form-driven ampliative inferences. In Saatsi's view, this distinction does more than just systematize the realism debate; it also makes apparent that the more form-driven arguments are seriously problematic. As a consequence, Saatsi suggests that realists should argue for what he calls 'rather content-driven' arguments. He sets out what he believes to be a challenge for any realist, namely to settle how best to construe these content-driven arguments in *general* terms.

Sherrilyn Roush focuses on the pessimistic meta-induction, which argues from the fact that past scientific theories have been proven false to the conclusion that our current theories will be proven false also. Roush suggests a novel solution to this argument, and argues that there is in fact no formulation of the argument that takes into account that the pessimist must appeal not merely to the falsity of our predecessors' theories but to the unreliability of their ways of coming to their beliefs (as confirmed by their repeated false conclusions). But even granting their unreliability, Roush argues that nothing follows from this about whether we have a right to our confidence in our particular theories. In establishing this, the pessimist must show that the reasons for believing that our predecessors were unreliable must be the same as the reasons for believing that we are unreliable but this has not been shown, since the difference in methods between us and our predecessors undermines the legitimacy of the pessimist's induction. Therefore there is no motivation for the suggested pessimism about the prospects for realism about science. As such, the ground has been cleared for the further development of realist positions in science.

2. Science and metaphysics

Metaphysics has had a somewhat rogue status amongst philosophers of science and the role of metaphysics continues to be a point of contention. With the collapse of logical positivism, no one would insist that metaphysical talk is *nonsense*. Nevertheless, the same empiricist sensibilities that motivate anti-realism motivate humble ontological commitments – metaphysics could be rejected as mere speculation.

Anjan Chakravartty asks about the proper role for metaphysical debate in philosophy of science and suggests that there is no one right answer. He is, in effect, arguing against the assumption that the debate between realists and anti-realists can be resolved. He is skeptical of the idea that there *should* be a single correct answer to the question of how much metaphysical speculation is appropriate in understanding science. Choices of where to draw the line, of how much metaphysics to do, are rather a matter of convention. Importantly, Chakravartty does not take this to motivate an empiricist stance about science; he takes himself to be a realist. His point is just that we should not expect there to be an Archimedean point from which one side can persuade the other to abandon their wrongheaded ways.

Jessica Pfeifer argues directly against one kind of anti-metaphysical view: nominalism. She argues that even the most sophisticated variety of nominalism will be unable to account for our inductive practices. A strict nominalist approach to language would leave us with no way of explaining that we do make inductive inferences, regardless of whether we are justified in doing so or not. In this way she exploits a fact about our epistemic practices, without taking a stance on whether this practice is justified. If her argument is successful, it's a victory for non-nominalist positions.

3. Representation

The question of how theories map the world is closely related to realism (the accuracy of those maps) and metaphysics (what the features are that appear on the maps). In the logical empiricist tradition, a theory was treated as a set of statements in a formal language. The question of how a theory represents was just the question of how language represents. Yet the statement view ultimately collapsed, and representation became a thornier problem. Chapters in this section address the general problem of scientific representation, the problem of determining when two representations are genuinely distinct (rather than just expressed differently), and the problem of visual representation.

Otávio Bueno sketches a framework for accommodating different features of scientific representation. Bueno views representation in scientific practice as being specific to contexts where only partial information is available. From this diversity we obtain partial isomorphisms between our theories and the objects represented. The framework promises to accommodate the diversity of representational mediums used in scientific practice, from models and templates to micrographs, and the output of various instruments.

Gregory Frost-Arnold and P. D. Magnus address the familiar problem of the underdetermination of theory by data – evidence is insufficient for establishing which of two different theories is correct. A case of would-be underdetermination could be averted if the alternatives were not *different* theories at all, but merely different formulations of the same underlying theory. They call this the identical rivals response, and argue that it involves reducing the ontological commitments of each formulation. They argue further that there is no logical criterion for determining whether this ontological withdrawal is appropriate. Instead, applying the identical rivals response (or not) is a strategic choice.

Laura Perini asks how visual images function as scientific representations. Since scientists use visual representations in their arguments, she insists, understanding scientific reasoning requires making sense of pictures. To make sense of them, she offers a semiotic analysis of pictures and diagrams as symbolic. This analysis also helps clarify scientific representation more generally.

4. The relation between science and values

It is a bit of hoary wisdom that science is *value free*. Philosophy of science is concerned with epistemic matters, and questions of value should be left to the ethicists. The chapters in this section explode this old conception.

Jay Odenbaugh considers what lessons general philosophy of science can learn by considering the environmental sciences. He focuses on three areas. First, he considers how environmental scientists rely on models and simulations which they know are inaccurate in important respects and uses this to inform our understanding of idealization in science generally. (This overlaps with the concerns about representation in the previous chapters.) Second, he considers the way value commitments affect environmental science and argues that it does not compromise scientific objectivity. Third, he considers the role that climate science plays in political debates. This role, he argues, raises questions about the function of controversy and consensus in science.

Justin Biddle and Eric Winsberg argue that, in the area of contemporary climate modeling, scientists cannot estimate the uncertainties of climate predictions in a way that is free from 'non-epistemic' considerations. To put the point differently, they argue that developments within climate modeling are influenced by values and choices. To be clear, they are not arguing that consensus regarding the causal connection between fossil fuel emissions and global climate change is problematic. Nor are they arguing for skepticism about the climate models that have been employed to reach this result. Rather, if they are right, the subtle influence of values in good climate science indicates that good science can be influenced by values.

Kristen Intemann's primary focus is feminist philosophy of science. Although concerned with how gender influences science, feminist work is more broadly concerned with how understanding this influence can inform our understanding of science more generally. For example, most feminist philosophers of science have held that social and political values can play a legitimate role in good science. Intemann distinguishes two positions: feminist empiricism and standpoint feminism. She argues that debates between these two positions are ultimately misguided. They are not as different as their debates might make them seem. Moreover, Intemann insists, a more plausible position can be developed by taking parts of each.

5. Science and social structure

Traditional philosophy of science has sometimes ignored the fact that science is a social activity. And it has also often taken the natural sciences, especially physics, to be paradigmatic of science. Social sciences were faced with two unsavory options: employ the methods of the natural sciences (naturalism) or articulate an alternative (interpretivism).

Daniel Steel argues that we should not accept either option in their traditional form. The methods of natural science cannot be characterized in a precise enough way to apply to both fundamental particles and societies. If we understand naturalism as the position that the natural and social sciences employ the same methods, then it is a non-starter. Yet this does not mean that interpretivists win all the traditional debates. Instead, Steel argues, we can see naturalism as a commitment to what he calls the Enlightenment ideal of a science. According to this ideal, social science aims to discover the causes and effects of social phenomena so as to inform social policy. Framing the debate in this way makes better sense in terms of general philosophy of science, and also promises to redirect the debate in more productive directions.

Michael Weisberg is concerned about science as a social activity, rather than with social science. He asks how the arrangement of scientists' cognitive labor affects the development of science – that is, how scientific communities might best be organized so as to facilitate progress and discovery. He discusses three areas of recent research: the *marginal contribution/reward* approach, the *epistemic networks* approach, and his own *epistemic landscape* approach. He discusses some specific work in the latter approach, modeling the scientific community as a mixture of followers and mavericks. He concludes by suggesting ways in which the approaches might inform one another.

Weisberg's chapter exemplifies our problem in trying to write a précis for this volume. The work he discusses has things to teach us both about specific questions and about science generally. Yet its importance for philosophy of science is not yet settled. And so it is with all the essays. In a decade or more, we may look back to see where the new waves crashed upon the old shore. But now we can only guess where each will end up.

We offer the present volume as a collection of things that seem important in the present philosophy of science, as judged by these authors.

9

Value Judgements and the Estimation of Uncertainty in Climate Modeling

*Justin Biddle and Eric Winsberg**

1. Introduction

It is uncontroversial that scientific research, especially scientific research that has important public policy implications, involves value judgements. This is nowhere more evident than in research on the impact of carbon emissions upon global climate. The looming prospect of severe, anthropogenic climate change is forcing us to make difficult moral and political decisions. What kind of action should be taken to curb global climate change? How much should we value our own safety, comforts, and economic opportunities in comparison to those of future generations? How much scientific evidence do we need before taking action? Should this action be voluntary or legally mandated? The importance of these questions for our future is well known, although there is still much disagreement over the appropriate way to answer them.

Yet, there are some questions regarding the role of value judgements in climate research that are not so well known, even within the climate modeling community. In particular, while it is clear that value judgements play an important role in deciding how a given area of research should lead us to *act*, it is less clear whether such value judgements should play a role in deciding what to *believe*. In other words, while value judgements clearly play a legitimate role in the realm of *practice*, do they also play a legitimate role in the realm of *theory*?

Philosophers of science are increasingly concerned with issues such as this, although, not surprisingly, different philosophers treat these issues in very different ways. The traditional view maintains, first of all, that the following two distinctions can be drawn clearly: the distinction between theory and practice and the distinction between so-called 'epistemic' and 'non-epistemic' values. It then maintains that only 'epistemic' values

play a legitimate role within the realm of theory; 'non-epistemic' values can, and should be, confined to the realm of practice.[1] According to one influential interpretation, 'epistemic values' – which include such values as simplicity, explanatory power, internal consistency, and consistency with surrounding theories – are values that are truth-conducive, in the sense that if theory T_1 exhibits a given 'epistemic' value and theory T_2 does not, then T_1 is, *ceteris paribus*, more likely than T_2 to be true; non-epistemic values, conversely, are not truth-conducive, and thus should be excluded from the realm of theory (McMullin, 1983).[2]

At various points during the last century, this traditional view has been called into question. In the early and mid-twentieth century, scholars such as C. West Churchman (1948, 1956), John Dewey (1929), Philip Frank (1954), Otto Neurath (1913), and Richard Rudner (1953) argued that values traditionally thought of as non-epistemic – including, in some cases, ethical and political values – play an inevitable role in the epistemic evaluation of research. Increasing numbers of contemporary philosophers of science are arguing for this same conclusion.[3] Despite the work of those who question this ideal of value-neutrality, however, the traditional view remains the dominant one.

In this chapter, we investigate the question of whether it is reasonable to expect climate modelers to exclude 'non-epistemic' values from the 'internal' aspects of their research – i.e. from the realm of theory. We argue that it is not. We begin our argument, in section 2, by discussing one of the most influential arguments for the claim that ethical values play an ineliminable role in the evaluation of scientific research, namely Richard Rudner's argument from inductive risk (Rudner, 1953); we follow this with a discussion of one of the most influential objections to this view, due to Richard Jeffrey (1956). Jeffrey's argument is the *locus classicus* of the view that one can distinguish clearly between the epistemic and the practical appraisal of theories and that the epistemic appraisal of research can and should be neutral with respect to 'non-epistemic' value judgements. Jeffrey's argument, moreover, is followed by many contemporary philosophers of science.[4]

We believe that an investigation of climate modeling is a particularly fruitful way to test the feasibility of Jeffrey's argument, because climate modelers, together with statisticians, attempt to do something very similar to what Jeffrey recommends – namely, to assign probabilities to hypotheses concerning the effects of carbon emissions upon global climate change in a manner that is free from 'non-epistemic' considerations. In section 3, we distinguish between the various sources of uncertainty in the development of climate models, and we discuss in general terms the ways in which climate scientists attempt to quantify these uncertainties. In sections 4 and 5, we discuss in detail two of these sources of uncertainty – structural model uncertainty and parameter uncertainty, respectively – and we argue that it is unreasonable to expect climate scientists to provide estimations of these uncertainties in a manner that is free from practical, or 'non-epistemic', considerations.[5]

Before we proceed to our discussion of Rudner and Jeffrey, there is one preliminary point that needs to be made right from the start. There is a broad consensus within the scientific community that carbon emissions are causally related to global climate change; we are convinced that this broad consensus is justified, and it is in no way our aim to call this consensus into question. We believe this consensus has been reached objectively. As will become apparent, we do not believe that the influence of 'non-epistemic' considerations on the estimation of uncertainties implies that climate models are unreliable. What it does imply is that, *pace* Jeffrey, one cannot distinguish sharply between the realms of value-neutral theory and value-laden practice. Moreover, it highlights the fact that more attention needs to be paid to the areas within climate modeling in which values play an ineliminable role, to the kinds of values or practical considerations that play a role, and to the effect that these values have upon the overall performance of our models. None of this, however, should be taken as evidence for skepticism about climate change. With this preliminary note aside, we can now proceed to our discussion of Rudner and Jeffrey.

2. Rudner and Jeffrey on the role of ethical values in science

In an essay entitled 'The Scientist *Qua* Scientist Makes Value-Judgments' Rudner argues that ethical values play an inevitable role in the epistemic appraisal of hypotheses. His argument proceeds as follows:

1. The scientist *qua* scientist accepts or rejects hypotheses.
2. No scientific hypothesis is ever confirmed with certainty. In accepting or rejecting a hypothesis, there is always the possibility of being wrong.
3. The decision to accept or reject a hypothesis depends upon whether the evidence is sufficiently strong.
4. Whether the evidence is *sufficiently* strong is 'a function of the *importance*, in a typically ethical sense, of making a mistake in accepting or rejecting the hypothesis' (Rudner, 1953: 2, his emphasis).
5. Therefore, the scientist *qua* scientist makes value judgements.

To illustrate this argument, Rudner considers the evaluation of the following two hypotheses: (a) a given drug, which we know contains a toxic ingredient that is lethal to human beings, is safe for human consumption, because it does not contain this ingredient in dangerous quantities; (b) a given batch of machine stamped belt buckles is not defective. Rudner argues that we require a much higher standard for accepting the former hypothesis than the latter, because the moral consequences of wrongly accepting the first hypothesis are much more serious than the second. 'How sure we need to be before we accept a hypothesis will depend on how serious a mistake would be' (Rudner, 1953: 2). Given that the degree of confirmation that we require in order to accept (or reject) a given hypothesis depends upon an evaluative judgement regarding potential ethical consequences, ethical considerations play an inevitable role in the appraisal of hypotheses. It should be noted that, in drawing this conclusion, Rudner is maintaining that value judgements, *even in the ideal*, play a role in the appraisal of hypotheses. It is not only the scientist *qua* human being – i.e. *qua* individual who is invariably influenced by the prejudices of her time – who makes value judgements; it is also the scientist *qua* scientist.

While Rudner does not spend much time examining the implications of this conclusion for a broader theory of science, he does discuss briefly its implications for the notion of scientific objectivity. According to a traditional interpretation, scientific objectivity requires that hypotheses be evaluated in a value-neutral fashion – or at least in a fashion that is neutral with respect to ethical values. Rudner does not deny that objectivity is and should be an ideal for scientific inquiry, but he argues that the necessarily value-laden character of hypothesis appraisal implies that this traditional interpretation of objectivity is misguided: 'What seems called for . . . is nothing less than a radical reworking of the ideal of scientific objectivity' (Rudner, 1953: 6). Rudner does not undertake the task of reworking this ideal; he does, however, argue that objectivity in science 'lies at least in becoming precise about what value judgments are being and might have been made in a given inquiry' (Rudner, 1953: 6). In other words, if value judgements do play an inevitable role in scientific research, it would be in the interests of objectivity to know precisely which value judgements are playing a role, where they are playing a role, and the effect that these judgements have upon the research in question.[6]

Many, however, deny that value judgements of an ethical – or more broadly, 'non-epistemic' – character play an inevitable role in the epistemic appraisal of hypotheses; one of the clearest arguments for this view was put forward by Jeffrey, in direct response to Rudner (Jeffrey, 1956). Jeffrey argues that the scientist *qua* scientist does not accept or reject hypotheses but merely assigns probabilities to them: 'the scientist's proper role is to provide the rational agents in the society which he represents with probabilities for the hypotheses which on the other account he simply accepts or rejects' (Jeffrey, 1956: 245). Once the scientist has assigned probabilities to hypotheses and communicated this information to 'rational agents' in society, these agents then assign utilities to the possible outcomes of accepting/rejecting the hypotheses and determine on the basis of a decision-theoretic calculation whether the hypothesis in question should be acted upon.

The primary argument that Jeffrey provides for the view that the scientist *qua* scientist does *not* accept or reject hypotheses is that the acceptance or rejection of a hypothesis *per se*, independent of a given practical context, is incoherent. In support of this, Jeffrey considers the following two hypotheses: (a) an entire lot of vaccine is safe, and (b) an entire lot of roller skate ball bearings is safe. In certain contexts, he argues, it is legitimate to demand a higher standard for the 'acceptance' of (a) than for (b) – but only when certain practical contexts are assumed, such as that the vaccine will be given to children. If we assume a different practical context, such as that the vaccine will be given to monkeys, the standards for 'acceptance' of the two hypotheses might be the same. Jeffrey summarizes his position by quoting approvingly the following passage from Bruno DeFinetti:

> I do not deem the usual expression 'to accept hypothesis H_r', to be proper. The decision does not really consist of this 'acceptance' but in *the choice of a definite action* A_r. The connection between the action A_r and the hypothesis H_r may be very strong, say 'the action A_r is that which we would choose if we knew that H_r was the true hypothesis'. Nevertheless, this connection cannot turn into an identification. (Quoted in Jeffrey, 1956: 242)

Thus, when one 'accepts' a hypothesis H, one is really choosing to act on the basis of H, and to do this requires that we specify a practical context in which the action is to occur. Because virtually all hypotheses could be acted upon in a multiplicity of different ways – e.g. the hypothesis that a given vaccine is safe could be used as a basis for vaccinating either children or monkeys – the notion of 'accepting' a hypothesis H *per se* is incoherent.

For our purposes, the primary implications of this view are (1) that one can distinguish clearly between the realms of theory and practice, or belief and action, and (2) that ethical considerations – or, in Jeffrey's terminology, considerations of utility – can be confined to the realm of practice. The scientist, Jeffrey argues, can and should remain in the realm of value-neutral theory and leave the ethical questions to 'the rational agents in the society which he represents'.

As stated earlier, Jeffrey's objection to Rudner's argument is still a standard objection to the view that ethical considerations play an inevitable role in the evaluation of research. For example, in response to Heather Douglas's argument that 'non-epistemic values are a required part of the internal aspects of scientific reasoning' (Douglas, 2000: 559), Sandra Mitchell maintains that this argument involves a 'conflation of the domains of belief and action [that] confuses rather than clarifies the appropriate role of values in scientific practice' (2004: 250). Moral and political values, according to Mitchell, play a legitimate role in the *practical*, not the *epistemic*, evaluation of research.

In the remainder of the chapter, we will argue that in the area of climate modeling, the Jeffreyan strategy does not succeed. In recent work, climate scientists, working alongside statisticians, have begun to attempt to do something very similar to what Jeffrey recommends; they attempt, that is, to estimate the uncertainties of various predictions made by climate models, and they attempt to do this in a manner that is free from moral, social, or any other kind of 'non-epistemic' value. They then hand these predictions and uncertainties over to policy-makers, legislators, and other representatives of the public, who are charged with determining how best to act.[7] In our view, however, this clean separation of the realms of value-neutral theory and value-laden practice is not realistically attainable. We will argue that one class of 'non-epistemic' values in particular – those that are reflected in decisions that certain types of prediction tasks are more important than others – influence the probabilities we assign to various possible climate outcomes. In order to argue for this, we need to establish a few preliminaries regarding the way in which uncertainties in climate modeling are estimated.

3. Climate modeling and uncertainty

Conceptually, we can distinguish three sources of uncertainty regarding the predictions of complex climate models. Firstly, there is uncertainty about the basic structure that our climate models ought to have. While the construction of climate models is guided by basic science, these models incorporate a barrage of auxiliary assumptions, approximations, and parameterizations, all of which contribute to a degree of uncertainty about the predictions of these models. We will call this type of uncertainty *structural model uncertainty*. Secondly, complex models involve large sets of parameters, or aspects of the model that have to be quantified before the model can be used to run a simulation of a climate system. We are often highly uncertain about what the best value for many of these parameters is, and hence, even if we had at our disposal a model with ideal (or perfect) structure, we would still be uncertain about the behavior of the real system we are modeling, because the same model structure will make different predictions for different values of the parameters. We will call uncertainty from this source *parameter uncertainty*. Finally, in evaluating a particular climate model, including both its structure and parameters, we compare the model's output to real data. Climate modelers, for example, often compare the outputs of their models to records of past climate. These records can come from actual meteorological observations or from proxy data – inferences about past climate drawn from such sources as tree rings and ice core samples. Both of these sources of data, however, are prone to error, and so we are uncertain about the precise nature of the past climate. This, in turn, has consequences for our knowledge of the future climate. We will call this source of uncertainty *data uncertainty*.

While data uncertainty is a significant source of uncertainty in climate modeling, we will not discuss this source of uncertainty here. For the purposes of this discussion, we will make the crude assumption that the data against which climate models are evaluated are known with certainty. We will be interested in arguing that values play an inevitable role in the estimation of uncertainties from the two other sources. How, then, do we estimate structural model uncertainty and parameter uncertainty? In both cases, there are, broadly speaking, two methods available to statisticians interested in quantifying these uncertainties. The first method makes use of *observable frequencies*, and the second *expert judgement*.

To understand the idea of using observable frequencies, consider the example of a simulation model with one parameter and several variables.[8] If one has a data set against which to benchmark the model, one could assign a weighted score to each value of the parameter based on how well it retrodicts values of the variables in the available data set. Based on this score, one could then assign a probability to each value of the parameter. Crudely speaking, what we are doing in an example like this is *observing* the *frequency* with which each value of the parameter is successful in replicating known data – how many of the variables does it get right? with how much accuracy? over what portion of the time history of the data set? – and then assigning this observed-frequency value to the probability of the parameter taking this value.

Of course, only in specific circumstances are frequencies probabilities. Absent other knowledge, it would naïve to think the observed frequencies in an example like the above are the actual probabilities of the values of those parameters. For one thing, we are interested in the best value of the parameter for predicting the behavior of the system for *all* times, not just the times for which we have sample data. For another, carrying out a procedure like the one above requires us to weight the relative importance of the various variables. Hence, while these kinds of frequencies can be useful guides in assigning probabilities to the values of a parameter, they are far from perfect, and some might view it to be more sensible to adopt a broadly subjectivist approach and to think that the best guide to these probabilities is the subjective degree of belief held by the best experts. Expert judgement, however, is not perfect either; in fact, expert judgement surely arises, *inter alia*, through the process of observing the degree to which model output matches available sample data. In practice, therefore, statisticians typically use some combination of expert judgement and observable frequencies to arrive at probabilities.

Thus, if we are interested in understanding where estimates of the degree of uncertainty about future climate come from, and, in particular, if we want to know to what degree these estimates are free from, or influenced by, various values, then we need to understand at least four things. We need to understand how observable frequencies and expert judgement are each used to estimate the degree of both structural model uncertainty and parameter uncertainty.

4. Structural model uncertainty

Let us begin with the estimation of structural model uncertainty. Given that one common method of estimating the degree of structural uncertainties in the predictions of climate models is by examining the degree of variation in the predictions of the existing set of climate models,[9] the range of models that happen to be available on the market will clearly influence our estimates of the degree of uncertainty contained in these models. In this section, we will argue that climate modelers have emphasized certain prediction and retrodiction tasks over others (e.g. predictions of global mean surface temperature change over predictions of global mean precipitation change), and that these past decisions have affected the performance of these models. We will then argue that this in turn invariably affects the current estimations of uncertainties. The decision to emphasize one prediction task over another is a paradigm example of a decision that reflects 'non-epistemic' values; that is, the decision is made not because emphasizing one prediction task over another has significant epistemic benefits, but rather because one set of prediction tasks is thought to be more important, in terms of its social, political, or economic consequences, than another. These value-laden decisions, we will argue, invariably affect the estimation of uncertainties of climate models.

As a preliminary to this argument, it is worth noting that the size of the dispersions of predictions of different climate models depends significantly upon the choice of a prediction task. Consider, for example, the graphs represented in Figure 9.1, which display predictions of temperature change and precipitation change made by different climate models.[10] For our purposes, what is important about these graphs is that the dispersion of predictions of mean precipitation change is significantly larger than the dispersion of predictions of global mean surface temperature change. The reason for this, we hypothesize, is that predictions of global mean surface temperature change have been more highly valued than other predictions; the climate modeling community has focused its energies upon refining and tweaking models in order to predict temperature change accurately, rather than precipitation change, sea level change, or any of a variety of other prediction tasks.

One might, however, question this hypothesis; after all, temperature and pressure are highly coupled variables in all of our models. Improvement in one should only be able to come in tandem with improvement in the other. Is there further evidence that the degree of structural model uncertainty that we get from 'observable frequencies' (from looking at the range of available models on the market) is affected by past values regarding the importance of various prediction tasks?

To answer this question, it is helpful to introduce what one might call the 'problem of attribution'. Since at least the late 1980s, the climate modeling community has been attempting to attribute the specific successes and failures of different climate models to specific components, or modules, of those models, for the purpose of developing models that predict all relevant quantities equally well. The thought was that once one understood the sources of disagreements of different climate models, one could take steps toward eliminating those disagreements. In 1989, a major program was founded at the Lawrence Livermore National Laboratory called the Program for Climate Model Diagnosis and Intercomparison (PCMDI), which had this as its stated goal. A number of projects were undertaken under the auspices of the PCMDI, including the Atmospheric Intercomparison Project, which began in 1990, and the Coupled Model Intercomparison Project, which began in 1995.[11]

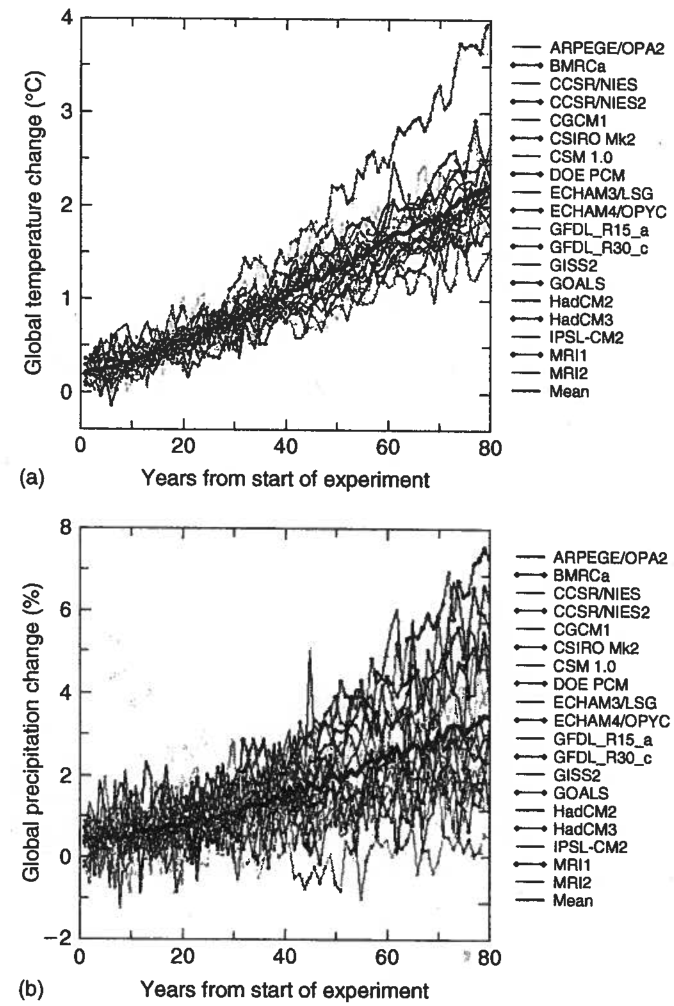


Figure 9.1 Graphs of global temperature change (°C) versus time (years) and global precipitation change (%) versus time (years)

Despite the optimism that surrounded the founding of the PCMDI, however, the project of attribution has been unsuccessful; not only this, there are strong reasons for believing that it will continue to be unsuccessful. Because the failure of the project of attribution is highly relevant to our argument that 'non-epistemic' values play an ineliminable role in the estimation of uncertainties, it will be necessary to discuss in some detail the reasons behind this failure.[12]

Kluges, generative entrenchment, and the historical character of model performance

The most sophisticated current climate models, Atmosphere-Ocean General Circulation Models (AOGCMs), are highly complex computer models that are constructed on the basis of both principled science – including fundamental partial differential equations from mechanics and thermodynamics – and trial and error approximations and parameterizations, and everything in between. An important feature of these models is their modular structure; each model is made up of components, or modules, and each module represents a given sub-system of the earth's climate, such as the circulation of the atmosphere, ice formation, ocean dynamics, cloud formation, the effects of vegetation, and the dynamics of aerosols. Today's climate models have their roots in the General Circulation Models of the 1950s, which described the atmosphere; over time, the complexity of these models has increased via the addition of more and more modules. The addition of these modules has allowed for better predictions of a growing range of phenomena; as already indicated, however, the predictions made by different climate models differ from one another, in some cases significantly.

An important reason for this variation in predictions is that the modules that make up a climate model are not mutually independent, but rather interact with one other, such that the overall performance of a model depends not simply upon the representational adequacy of each individual module, but also upon the way in which the modules are coupled together. Thus, if module X is good at predicting phenomenon Y, and if module W is good at predicting phenomenon Z, it will not, at least in general, be the case that the model consisting of X and W will be good at predicting both Y and Z. This overall model might be good at predicting one or the other of these phenomena, or neither, depending upon the way in which the modules interact with one another.

Moreover, not only does the overall performance of a given climate model depend upon the interaction of its modules; it is also the case that the specific ways in which modules interact with one another will depend upon the overall model of which they are 'parts'. The reason for this is that the process of coupling modules together relies to a large extent upon fitting the module to the overall model – not only to its principled structure but also its parameterizations.[13] This, in turn relies on

trial and error, piecemeal mutual adjustments of parameters and parameterization schemes – adjustments that need to be undertaken in different ways, depending upon the details of the model and modules in question. In order to illuminate the nature of these adjustments, a number of commentators have employed the notion of a 'kluge', a colloquial term first employed by computer programmers to describe sections of code that were functional but unprincipled, inelegant, and ill-understood. According to Andy Clark, a kluge is 'an inelegant, "botched together" piece of program; something functional but somehow messy and unsatisfying' (Clark, 1987: 278). The term describes aptly the process of coupling two modules together; while parts of the process might be principled, there are other parts that rely to a large extent upon simply botching things together.

In our view, the rather messy character of module coupling has important implications for the development of climate models; it suggests that model development is an historical process that depends in important respects upon the environment in which that model develops. To see this, it is helpful to draw upon William Wimsatt's notion of 'generative entrenchment' (Wimsatt, 2007). According to Wimsatt, 'a deeply generatively entrenched feature of a structure is one that has many other things depending on it because it has played a role in generating them' (2007: 133). Wimsatt employs this notion in order to explain the relationship between biological development and evolution; characteristics that are adaptive for one organism will not, in general, be adaptive for another, because the different organisms will, in general, have different features that are generatively entrenched. Unlike Dumbo, real elephants will never be able to fly, because they have particular features – e.g. bulkiness – that have developed in specific circumstances as a result of specific environmental pressures and that make adaptations such as wings impossible. Bulkiness is a generatively entrenched feature of elephants. Analogously, there are some climate models that will never be successful at predicting a particular phenomenon via a particular module, because the model has developed in specific circumstances as a result of specific 'environmental' pressures – e.g. pressures to emphasize certain predictive tasks over others – that make a given module impossible. A particular method of kludging the model might be as generatively entrenched as bulkiness is to an elephant. Just as tusks won't help a bird and wings won't help an elephant, a module that helps model A might not help model B.

Consider the following idealized example. Climate Model A is a simple General Circulation Model that describes the dynamics of the atmosphere. By building upon A, two further models, B and C, develop, both of which describe the following sub-systems of the earth's climate: atmosphere dynamics, ocean dynamics, ice formation, and the effects of vegetation. Model B develops from A via the following path: first, a module for ocean dynamics is coupled to Model A, followed by a module for ice formation, and finally a module for the effects of vegetation. Model C took a slightly different path; first, a module for ocean dynamics is coupled to A, followed by a module for the effects of vegetation, and finally a module for ice formation. Given the fact that B and C took different developmental paths, the attempt to couple additional modules to B and C – e.g. a module for aerosols – will, in general, need to proceed along different lines for each model; the piecemeal approximations to parameterization schemes that will allow a new module to be coupled to B will, in general, not work for C, and vice versa. This does not mean that it will be impossible to add an aerosol module to C; however, given the ways in which B and C have developed, it might be significantly more difficult to add this module to C than to B, and it almost certainly will need to be done via a very different coupling process. These difficulties, moreover, only increase with the complexity of the model in question. Thus, given models of the complexity that we have today, the historical development of these models – which, again, has been influenced significantly by the 'environmental' pressures placed upon them – places constraints upon the kinds of modules that can be coupled to them, which in turn affects the overall performance of the models.

The preceding discussion should make clear why the problem of attribution has been, and will in all likelihood continue to be, unsuccessful. One cannot attribute the predictive success or failure of a given model to a particular, localized component of that model because the components of models are strongly coupled to one another, and hence interact with one another in significant fashion. Moreover, the kinds of additions that can be made to a climate model – e.g. the ways in which the scope of a model can be expanded – will in general depend upon the way in which that model has developed over time, including the approximations and adjustments that have been made in order to couple new modules successfully. This, in turn, will depend upon the 'environmental' pressures to which the model is subject. All of this suggests that previous decisions to emphasize certain prediction and retrodiction tasks over others – e.g. predictions of global mean surface temperature change over predictions of global precipitation change – have a significant effect upon the ways in which these models can be expanded, and upon the ways that future expansions of these models will perform.

Values and the estimation of structural model uncertainty

What does all of this have to do with the debate over the role of values in scientific research? According to the preceding discussion, the overall performance of a model – its ability to predict and retrodict certain tasks well, and its inability to predict and retrodict other tasks well – depends to a significant extent upon the history of that model, which in turn depends upon that model's 'environment', which includes political decisions to emphasize certain predictive and retrodictive tasks over others. Because the estimation of uncertainties of our knowledge of climate change depends upon the performance of our best climate models, and because decisions of a 'non-epistemic' character have an ineliminable effect upon the performance of our best climate models, such 'non-epistemic' values invariably influence the estimation of uncertainties. Thus, Jeffrey's claim that we can assign probabilities to hypotheses in a value-neutral fashion – or, in the present case, that we can assign uncertainties to predictions in a value-neutral fashion – is false, at least in the area of climate modeling.

There are a number of ways that one might object to this argument. First of all, our argument depends upon the claim that the choice of prediction task is inevitably influenced by 'non-epistemic' factors; yet one might argue that the decision to emphasize temperature predictions over other predictions can be explained solely on theoretical, or 'purely epistemic', grounds. While there are a variety of ways in which one could develop this objection, perhaps the most plausible would proceed as follows. Of the candidate prediction tasks that one could choose to emphasize – including global mean surface temperature change, ocean heat content change, polar ice cap retreat, and so on – temperature is the most significant theoretically, because all of the other quantities can be derived from it. Without knowing how global mean surface temperature will change, one cannot explain or make sense of any other change in global climate; therefore, predictions of temperature change should be emphasized.

The objection thus formulated, however, has its problems. It is true that CO_2 and other gases are often called 'heat trapping', but it would be more accurate to call them 'energy trapping', and one of the most difficult problems within contemporary climate research is determining where this energy will go. Will it go towards heating the ocean? Melting sea ice? Or raising temperature? If it goes towards the latter, will it do this on the surface of the earth or in other parts of the atmosphere? To maintain that all of these other features of the planet's dynamics can be *derived* from mean global surface temperature is misguided, because it fails to take into account the massive and not entirely understood interdependencies that exist in the climate. All of these interdependencies together react in complex fashion to the trapping of additional energy, and it would be a mistake to assume that this additional energy will always manifest itself as higher surface temperatures. Increased surface temperature is only one of many symptoms of global climate change.

It seems to us that this is the most plausible way of developing this objection, and yet it still fails. One could, of course, attempt to formulate it in another fashion. We are skeptical, however, that such a route will be successful. Until this burden is met, therefore, we conclude that the decision to emphasize predictions of global mean surface temperature change is one that is ineliminably influenced by 'non-epistemic' considerations, such as the fact that this prediction task was seen as the most important, in terms of its social, political, and economic consequences.

A second objection that one might direct against our argument is that the way in which values operate in this case is philosophically uninteresting; if scientists decide, for whatever reason, to put more resources into the prediction of one task than another, it is hardly novel to maintain that the resulting models will be better at predicting the former task than the latter. This objection, however, fails for at least two reasons.[14]

First of all, the objection obtains its intuitive force via a false presupposition, namely that it is realistic to attempt to construct models that perform in a manner that is independent of the influence of value judgements. According to the previous discussion, the performance of current climate models depends *invariably* upon the historical development of those models, which in turn depends *invariably* upon decisions of a social or political character, including decisions that a given prediction and retrodiction task is more important, from a social, moral, or economic point of view, than another. If it were reasonable to expect the performance of climate models to be independent of such past judgements, then the objection would be a good one; according to our argument, however, this expectation is not at all reasonable.

Secondly, and perhaps more importantly, the objection misrepresents the effects of focusing upon temperature predictions at the expense of precipitation predictions. In particular, the objection suggests that the primary effect of focusing upon one kind of prediction task over another is a difference in the actual uncertainties associated with these two predictive tasks. Yet, it is not merely the case that a lack of focus upon precipitation predictions will result in models that perform poorly with respect to this prediction task – and thus to higher actual uncertainties about future precipitation; it is also the case that a focus on mean global surface temperature predictions at the expense of global precipitation predictions will likely result in a range of models that *overstates* our estimates of the uncertainty of precipitation predictions and *understates* the uncertainty of temperature predictions. This choice of prediction tasks, in other words, affects not only the actual uncertainties of the models in question, but also the *estimations* of these uncertainties.

The reason for this pertains to the problems associated with estimating structural model uncertainty via observable frequencies – or, more specifically, via observing the range of predictions generated by our existing arsenal of models. If there are some particularly bad models on the market, we will overestimate the degree of uncertainty that we ought to have about our best models. On the other hand, if there is some degree of co-evolution in our models – that is, if the makers of our models deliberately ensure that the predictions of their model do not deviate too much from the herd (a natural thing to do if you do not want it to seem like the predictions of your model are unrealistic) – then the range of models on the market will underestimate the degree of uncertainty that we ought to attribute to them. As Myles Allen, who refers to this method as using 'ensembles of opportunity', notes: 'If modelling groups, either consciously or by "natural selection", are tuning their flagship models to fit the same observations, spread of predictions becomes meaningless: eventually they will all converge to a delta-function' (Allen, 2008). Thus, by misrepresenting the effects of focusing upon temperature predictions at the expense of precipitation predictions, the objection obscures the significant impact that values have in this situation.

Structural model uncertainty and expert judgement

Of the four modalities that we discuss in this chapter (expert judgement versus observable frequencies and structural model uncertainty versus parameter uncertainty), we have thus far restricted our attention to the estimation of structural model uncertainty via observable frequencies. The aforementioned problems with this method, however, lead to a further challenge to our claim that 'non-epistemic' considerations play an ineliminable role in the estimation of structural model uncertainty – namely that one could perhaps make such estimations via expert judgement, rather than observable frequencies, and do so in a manner that is free from 'non-epistemic' values. Are there reasons for believing that values play an ineliminable role in the case of expert judgement as well?

While we believe that there are such reasons, the question is nevertheless rather difficult to answer, primarily because, of the four modalities under consideration, the use of expert judgement to estimate structural model uncertainty is the one with the least developed methodology in the literature.[15] The source of this lacuna is clear; when statisticians solicit the opinions of experts regarding the degree of structural uncertainty in a model, the statisticians are often confronted with the problem that the experts are either unwilling or unable to assign subjective degrees of belief to the structure of a model. In particular, modelers are extremely reluctant to assign probabilities to the predictions of models that are deterministic, given a particular value in parameter space.[16] When statisticians, trying to implement this modality, make it clear that what they are after is the modelers' subjective degree of belief, the modelers are apt to balk. Thus, it is not at all clear that the relevant experts actually possess the subjective degrees of belief, and methods need to be created that will reliably elicit these judgements.

The issue that is of interest to us, of course, is the extent to which experts' judgements about their subjective degrees of belief, assuming that they are willing to make these judgements, are influenced by past decisions about prediction tasks. Given, again, that a well-established methodology for making these judgements does not yet exist, the issue is a difficult one to resolve. We can, however, make the following preliminary argument. In order for expert judgement to overcome the effects that past prediction task priorities have had on ensembles of existing models, the experts in question will have to understand exactly what these effects have been. Yet, the failure of the project of attribution, discussed above, leads us to be skeptical that one can know what the precise effects of past prioritizations upon existing models actually have been. This, in turn, makes us skeptical that expert judgement can be free from 'non-epistemic' considerations, such as the choice of one prediction task over another.

Where does this leave us? We have argued that estimations of structural model uncertainty via observable frequencies cannot be done in a value-free way, and we have called into question the claim that this can be done in the case of expert judgement as well. In order to establish this latter point more firmly, more research on the methodology of this modality needs to be done. In the meantime, we have to be content with the fact that the predominant method in actual use in the scientific community is to look at observable frequencies; this method, as we have argued, is not value free.

Suppose, however, that one is unconvinced by the arguments just provided. Even if one continues to maintain that the estimation of structural model uncertainty can be done independently of 'non-epistemic' considerations – indeed, even if one assumes that there is no uncertainty whatsoever about the structure of a model – there is still significant uncertainty about the values of parameters and parameterization schemes. We believe that there is an even stronger case to be made for the claim that 'non-epistemic' considerations affect the estimation of parameter uncertainty. It is to this argument that we will now turn.

5. Parameter uncertainty

To keep the issue of parameter uncertainty conceptually distinct from that of structural model uncertainty, let us suppose that we know that we have a structurally perfect model with n parameters and that we are uncertain about what the best value is for each of these parameters. We can think of the set of n parameters as forming an n-dimensional parameter space. If we had a probability density function (PDF) over that n-dimensional space, we could easily assign probabilities to the predictions of the model. In principle, we would use something like the following Monte Carlo method: we could sample from the space of parameters in accordance with the given PDF, and calculate the output of the model for each of the sampled points in the space.[17] The resulting set of outputs would have means and variances for each of the variables that would correspond to the uncertainties we would assign to the corresponding predictions. If, for example, the mean prediction of the probability weighted sample for temperature is x, with standard deviation of s, then we could say that it is 95 percent likely that the modeled system's temperature will be $x \pm 2s$.

The question, then, is where a PDF over the space of possible parameter values can come from. The answer, as noted, is from observed frequencies and expert judgement. We've already seen, in some detail, how the former works. Each value in parameter space can be benchmarked; it can be scored for its ability, when used as input for a given model structure, to reproduce existing data, and probability densities can be assigned in proportion to that score. It is important to remember, however, that climate models are highly multivariate in their output. Scoring the performance of each point in parameter space, therefore, requires us to weight the relative importance of making good predictions of each variable. Such a weighting involves a value decision that involves both 'epistemic' and 'non-epistemic' considerations.[18] More importantly, expert judgement might very well affirm that a point in parameter space that is very successful at reproducing known, available data will be poor at predicting unknown, unavailable data. In practice, therefore, a more common procedure is to use a benchmarking procedure to rule out certain regions of parameter space – if a point in parameter space does a very poor job of reproducing known data then it can be thrown out – and then use expert judgement to assign a PDF to the remaining subset of the space.

But what are experts judging when they assign such a PDF? Here, we think it is important to keep in mind an important difference between two kinds of parameters that climate models can take as inputs. So far, we have simply defined a parameter of a model as any aspect of the model that has to be quantified before the model simulator can be run. But the aspect of the model being quantified might or might not correspond to some actual quantifiable or measurable property of the physical system being modeled. While some of the parameters of our models are of a kind with physical parameters like g, the acceleration due to gravity near the surface of the earth ($9.8\,m/s^2$), others of them, which result from so-called 'parameterizations' in the model, are rather different.

Parameterizations are elements of a simulation model that are designed to capture effects that slip between the cracks of a model's discretization grid or are otherwise lost to an approximation of the model.

> Sub-grid processes are represented by *parameterisation schemes* describing their aggregated effect over a larger scale. These schemes are often referred to as 'model physics' but are really based on physics-inspired statistical models describing the mean quantity in the grid box, given relevant input parameters. (Benestad, 2007)

Parameterization schemes are extremely common in global climate models, especially given that reductions in discretizations and approximations are bought at the price of complexity of computer models, increase of simulation time, and so on (ibid.).

In our view, there is an important difference between these two different kinds of parameters. With respect to the first kind of parameter, it makes sense to talk about the correct value of the parameter. For example, near the surface of the earth, the correct value for g is approximately $9.8\,m/s^2$. But the value of a parameter associated with a parameterization scheme does not have a single correct value. At best, it has a best value for a particular prediction task – a value which, if used as model input, will enable the model to make the best possible predictions for that particular task.

Goldstein and Rougier highlight the important difference between these two kinds of parameters when it comes to estimating uncertainties.

In their discussion of eliciting expert judgement about appropriate PDFs to place on parameter space, they write:

> In practice, modellers often seem to take two somewhat contradictory positions about the status of the simulator's best input, on the one hand arguing that it is a hypothetical construct and on the other hand using knowledge and intuition derived from the physical system to set plausible intervals within which such a value should lie. (Goldstein and Rougier, 2006: 5)

Of course, when it comes to parameterization schemes, there is no straightforward relation between the values of the inputs to the simulator and any measurable corresponding physical values for the system. If expert judgement is to be relied upon in assigning probabilities to parameter values, therefore, it has to come not from *knowledge and intuition of the physical system alone*, but from knowledge and intuition of the behavior of the model vis-à-vis the system. In that regard, one more passage from the Goldstein and Rougier piece is particularly striking: 'In particular, there was strong disagreement [among statistically numerate system experts] with [the following:] that there exists an input x* such that, were it to be known, only a single evaluation of the simulator would be necessary' to learn everything that the model has to tell us about the system (ibid.). What the statistically numerate system experts are suggesting is that the best value from parameter space for one particular prediction task is not necessarily the same as the one for another prediction task. Indeed, it may even be that for some prediction task, the maximum amount of knowledge that can be extracted from a particular model structure might come from an ensemble of runs from multiple parameter values. Thus, the idea that there is a single 'best value' for a parameter in a model is, in many cases, not correct. The best value for prediction task A might not be the same as for prediction task B. And even for a single prediction task, it might be an ensemble of different parameter values that yields the most valuable knowledge.

What is important to note here is that, as emphasized in the previous section, climate models can be applied to a variety of prediction tasks. While climate modelers have traditionally emphasized predictions of global mean surface temperature change, they have more recently begun to turn their attention to other prediction tasks, such as the probability of so-called 'abrupt' climate changes, including the collapse of thermohaline circulation, the drying of the Amazon, and the disappearance of Arctic summer sea ice. What we do not know for a particular climate model with all of its parameterizations, is whether the best parameter value for predicting mean surface temperature is also the best parameter value for predicting, say, the extent of thermohaline circulation. But if the claim made by Rougier and Goldstein's statistically numerate system experts is right (and we think it is), and they are not always the same, then the following suspicion arises.

Suppose we are confronted by a novel prediction task, such as estimating the probability of thermohaline collapse. And suppose that, for a particular model structure, we want to estimate the degree of parameter uncertainty about such predictions. To do this, as we have discussed, would require us to use expert judgement to assign a PDF to the parameter space for that model. If what experts are judging is where, in parameter space, the 'best' value of parameters can be found, and if the best value for predicting mean surface temperature is not necessarily the best value for predicting thermohaline circulation, then the suspicion is that expert judgement about parameter space PDFs will be influenced, to some degree or other, by the intended purposes for which these experts have been constructing and testing their models during the period in which they have acquired their expertise. Again, when it comes to judgement about parameterizations, the relevant expertise is not about the system itself, but about the behavior of the model vis-à-vis the system.

Values and the estimation of parameter uncertainty

At this point, we are in a position to state the argument for the ineliminability of 'non-epistemic' considerations in the estimation of parameter uncertainty. Suppose that we have a particular model structure, about which there is no uncertainty, and suppose furthermore that we want to employ this model in order to estimate the probability of thermohaline collapse given a doubling of atmospheric carbon dioxide. To estimate this, we have to elicit expert opinion regarding an appropriate PDF to assign to the space of parameters. But the best value(s) of parameters for predicting thermohaline circulation is not necessarily the same as the best value(s) for other prediction tasks, such as predicting mean surface temperature. If the experts whose opinion we elicit have predominantly acquired experience with their models in predicting mean surface temperature, then this will affect their judgement about where the best value for parameters can be found.

But the particular set of prediction tasks that have played a role in shaping our experts' judgements have been the product of a set of choices – for example, the choice to focus on predicting mean surface temperature rather than mean global precipitation. And these choices, in turn, reflect

a set of values – namely the set of social, economic, or other considerations that have historically led us to believe that predicting temperature is more important than predicting precipitation. These values, in other words, are ones that are traditionally regarded as being 'non-epistemic' in character. Thus, 'non-epistemic' values play an ineliminable role in the estimation of parameter uncertainty.

6. Conclusion

Jeffrey argues that the task of the scientist can and should be limited to the assignment of probabilities to hypotheses, and that this task can be carried out in a manner that is free from practical, 'non-epistemic' considerations. Once the task of assigning probabilities to hypotheses has been completed, the scientist can then hand these hypotheses and probabilities over to policy-makers, who are responsible for deciding upon a course of action. On this picture, there is a clean separation between the realms of theory and practice and a clear line that divides the space where values play a legitimate role (the realm of practice) and where they do not (the realm of theory). As noted earlier, Jeffrey's line of reasoning is still very commonly followed today.

If our argument is sound, however, Jeffrey's line of reasoning fails, at least in one very important area of contemporary scientific research – namely, climate modeling. Scientists cannot assign probabilities to hypotheses about climate change – or, more specifically, estimate the uncertainties of climate predictions – in a manner that is free from 'non-epistemic' considerations, because 'non-epistemic' considerations invariably influence the choices of prediction tasks, and the choices of prediction tasks invariably influence the estimation of both structural model uncertainty and parameter uncertainty.

Again, we do not believe that this result in any way implies that the consensus that has been formed regarding the causal connection between fossil fuel emissions and global climate change is problematic. We are not skeptics of climate change. We do, however, believe that this conclusion suggests that more attention should be paid to the spaces within climate modeling where values play a role, to the kinds of values or 'non-epistemic' considerations that play a role, and to the effects that these values have upon the overall performance of our models.

Notes

*We have accrued many debts in writing this chapter. Thanks to Heather Douglas and Elizabeth Lloyd for their comments on an earlier draft. Thanks to James

McAlister, Michael Weisberg, Ron Giere, and others for their comments on presented versions of this paper in Dubrovnik and Mexico City. We began thinking about this topic as Fellows of the research group, Science in the Context of Application, at the Center for Interdisciplinary Research (ZiF) at Bielefeld University. Thanks to Martin Carrier and Alfred Nordmann for organizing this group, and thanks to Johannes Lenhard, Torsten Wilholt, Cornelis Menke, and Justus Lentsch for listening to some early ideas. Special thanks to Johannes for allowing us to borrow liberally from his and Eric's collaboration. Eric would like to thank the IAS, Durham University, for giving him time to work on this chapter, and for exposing him to the expertise of several statisticians who work on climate models, especially Michael Goldstein, Peter Challenor, and John Haslett.

1. Defenders of this view include Giere (2003), Jeffrey (1956), Kitcher (2001), Koertge (2003), McMullin (1983), and Mitchell (2004).
2. It has not always been standard to allow a role for values of any kind in the appraisal of theories. One of the primary aims of the neo-positivist movement in post-Second World War America was to explicate fundamental methodological concepts such as confirmation and explanation in purely formal terms (e.g. Hempel, 1945; Hempel and Oppenheim, 1948). Had this project succeeded, there would be no need for any values – including epistemic values – in the appraisal of theories. The project, however, did not succeed, and since the demise of neo-positivism, it has been standard to argue that values of some sort play an inevitable role in theory appraisal. For classic expressions of this view, see Kuhn (1969, 1977) and McMullin (1983).
3. See, for example, Douglas (2000), Howard (2006), Longino (1990, 2002), Kourany (2003a, 2003b), Solomon (2001), and Wilholt (forthcoming).
4. For example, Ron Giere employs a version of this objection in his criticism of Janet Kourany's proposal for a socially responsible philosophy of science (Giere, 2003; Kourany, 2003a, 2003b), and Sandra Mitchell employs it against the arguments of Heather Douglas (Douglas, 2000, 2004a, Mitchell, 2004).
5. Throughout this chapter, we will employ the traditional terminology of 'epistemic' and 'non-epistemic' values. We recognize, however, that this terminology is in many respects problematic, as has been argued in Longino (1996), Rooney (1992), and a number of the works cited in note 3. As is apparent from our previous discussion, the terminology arose in the context of arguments to the effect that social values, moral values, and other broadly practical considerations can and should be excluded from the epistemic evaluation of research – hence the label 'non-epistemic'. If it turns out, however, that such values do play a legitimate role in the epistemic evaluation of research, then it makes little sense to label them 'non-epistemic' values. Thus, our essay – in addition to the literature cited in note 3 – provides strong reasons for abandoning the terminology of 'epistemic' and 'non-epistemic' values. However, given the entrenched character of this terminology, we have decided in the interests of convenience to use it here.
6. See Douglas (2004a), Fine (1998), and Longino (1990) for discussions of the compatibility of values and objectivity in science.
7. For a clear example of this, see the so-called 'Stern Review' (Stern, 2007).
8. A parameter for a model is an input that is fixed for all time, while a variable takes a value that varies with time. A variable for a model is thus both an input for the model (the value the variable takes at some initial time) and an output (the value the variable takes at all subsequent times). A parameter is simply an input.

9. See, for example, IPCC (2007) or, for a more skeptical position regarding the feasibility of this endeavor, Smith (2002). The typical method is simply to take the average and standard deviation of the existing predictions and generate uncertainties from these values.
10. Taken from IPCC (2001: 537).
11. For further details, see the PCMDI website: http://www-pcmdi.llnl.gov/about/index.php (accessed May 13, 2008).
12. The following discussion draws heavily upon Winsberg and Lenhard (forthcoming).
13. Parameterizations, roughly speaking, are correction factors, or attempts to correct for the fact that simulation models cannot capture effects that occur at scales below which the models are discretized. We discuss parameterizations in much greater detail in section 5.
14. Thanks to James McAllister for encouraging us to think further about this objection.
15. For a discussion of these methods, see Goldstein and Rougier (2006). For criticism of these methods, see Allen (2008).
16. Modelers, in fact, often think statisticians will do that for them. Statisticians, on the other hand, clearly (and correctly) view this as a scientific problem whose answer can only be provided by those with the best scientific expertise.
17. We say 'in principle' because in practice the situation is more complicated. In practice, it is too time consuming to calculate the outputs of a global climate model for sufficiently many sampled points. Statisticians, therefore, use so-called 'emulators' of the models, in effect simulations of the simulations, to calculate these outputs. While this introduces a large layer of technical difficulties to the project, it has no bearing, that we are aware of, on what we discuss here.
18. Here, we think, is an interesting example of where the purported distinction between epistemic values and non-epistemic values breaks down. Whether or not the value underlying such a weighting would count as epistemic or non-epistemic would depend on whether the researchers could argue that their choice of weighting was made on the basis of which weighting was the 'best guide to truth'. But which truths? And how would such an argument be resolved? The distinction here becomes a bit confused.

References

Allen, M. (2008) 'What Can be Said about Future Climate?' http://www.climateprediction.net/science/pubs/allen_Harvard2008.ppt (accessed July 3, 2008).
Benestad, R. (2007) 'Why Global Climate Models Do Not Give a Realistic Description of the Local Climate', *RealClimate*, http://www.realclimate.org/index.php/archives/2007/05/climate-models- local-climate (accessed May 29, 2007).
Churchman, C. West (1948) *Theory of Experimental Inference* (New York: Macmillan).
Churchman, C. West (1956) 'Science and Decision Making', *Philosophy of Science*, 22: 247–249.

Clark, A. (1987) 'The Kluge in the Machine', *Mind and Language*, 2: 277–300.

Dewey, J. (1929) *The Quest for Certainty* (New York: Minton, Balch & Company). Reprinted in J. A. Boydston (ed.), *John Dewey: the Later Works, 1925–1953*, vol. IV (Carbondale and Edwardsville: Southern Illinois Press, 1988).

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

Douglas, H. (2004a) 'Border Skirmishes between Science and Policy: Autonomy, Responsibility, and Values', in P. Machamer and G. Wolters (eds), *Science, Values, and Objectivity* (Pittsburgh: University of Pittsburgh Press), 220–244.

Douglas, H. (2004b) 'The Irreducible Complexity of Objectivity', *Synthese*, 138: 453–473.

Fine, A. (1998) 'The Viewpoint of No-One in Particular', *Proceedings and Addresses of the American Philosophical Association*, 72: 9–20.

Frank, P. G. (1954) 'The Variety of Reasons for the Acceptance of Scientific Theories', in P. G. Frank (ed.), *The Validation of Scientific Theories* (Boston: Beacon Press), 3–17.

Giere, R. N. (2003) 'A New Program for Philosophy of Science?', *Philosophy of Science*, 70: 15–21.

Goldstein, M. and J. Rougier (2006) 'Reified Bayesian Modelling and Inference for Physical Systems', reprint for *Journal of Statistical Planning and Inference*.

Hempel, C. G. (1945) 'Studies in the Logic of Confirmation', *Mind*, 54: 1–26, 97–121. Reprinted in Hempel (1965), 3–46, with a 1964 Postscript added.

Hempel, C. G. (1965) *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: The Free Press).

Hempel, C. G. and P. Oppenheim (1948) 'Studies in the Logic of Explanation', *Philosophy of Science*, 15: 135–175. Reprinted in Hempel (1965), 245–290, with a 1964 Postscript added.

Howard, D. A. (2006) 'Lost Wanderers in the Forest of Knowledge: Some Thoughts on the Discovery–Justification Distinction', in J. Schickore and F. Steinle (eds), *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction* (New York: Springer).

IPCC (2001) *Climate Change 2001: the Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press).

IPCC (2007) *Climate Change 2007: the Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press).

Jeffrey, R. C. (1956) 'Valuation and Acceptance of Scientific Hypotheses', *Philosophy of Science*, 22: 237–246.

Kitcher, P. (1993) *The Advancement of Science: Science without Legend, Objectivity without Illusions* (New York: Oxford University Press).

Kitcher, P. (2001) *Science, Truth, and Democracy* (New York: Oxford University Press).

Koertge, N. (2003) 'Screen Out Social Values from Core Epistemic Areas of Science', in C. Pinnick, R. Almeder, and N. Koertge (eds), *Scrutinizing Feminist Epistemology: an Examination of Gender in Science* (New Brunswick, NJ: Rutgers University Press).

Kourany, J. (2003a) 'A Philosophy of Science for the Twenty-First Century', *Philosophy of Science*, 70: 1–14.

Kourany, J. (2003b) 'Reply to Giere', *Philosophy of Science*, 70: 22–26.

Kuhn, T. (1962) *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press).

Kuhn, T. (1969) 'Postscript – 1969', in *The Structure of Scientific Revolutions*, 2nd edn (Chicago: University of Chicago Press), 174–210.

Kuhn, T. (1977) 'Objectivity, Value Judgment, and Theory Choice', in *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Chicago: University of Chicago Press), 320–339.

Longino, H. (1990) *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry* (Princeton, NJ: Princeton University Press).

Longino, H. (1996) 'Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy', in L. H. Nelson and J. Nelson (eds), *Feminism, Science, and the Philosophy of Science* (Dordrecht: Kluwer Academic Publishers), 39–58.

Longino, H. (2002) *The Fate of Knowledge* (Princeton, NJ: Princeton University Press).

Machamer, P. and G. Wolters (eds) (2004) *Science, Values, and Objectivity* (Pittsburgh: University of Pittsburgh Press).

McMullin, E. (1983) 'Values in Science', in P. D. Asquith and T. Nickles (eds), *Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association*, vol. I (East Lansing, MI: Philosophy of Science Association), 3–28.

Mitchell, S. D. (2004) 'The Prescribed and Proscribed Values in Science Policy', in P. Machamer and G. Wolters (eds), *Science, Values, and Objectivity* (Pittsburgh: University of Pittsburgh Press), 245–255.

Neurath, O. (1913). 'Die Verirrten des Cartesius und das Auxiliarmotiv. Zur Psychologie des Entschlusses', in *Jahrbuch der Philosophischen Gesellschaft an der Universität Wien* (Leipzig: Johann Ambrosius Barth). English translation: 'The Lost Wanderers of Descartes and the Auxiliary Motive (On the Psychology of Decision)', in Otto Neurath, *Philosophical Papers 1913–1946*, ed. and trans. R. S. Cohen and M. Neurath (Dordrecht: D. Reidel), 1–12.

Polanyi, M. (1962) 'The Republic of Science: Its Political and Economic Theory', *Minerva*, 1: 54–74.

Rooney, P. (1992) 'On Values in Science: Is the Epistemic/Non-Epistemic Distinction Useful?', in D. Hull, M. Forbes, and K. Okruhlick (eds), *Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association*, vol. II (East Lansing, MI: Philosophy of Science Association), 13–22.

Rudner, R. (1953) 'The Scientist *Qua* Scientist Makes Value Judgments', *Philosophy of Science*, 20: 1–6.

Smith, L. A. (2002) 'What Might We Learn from Climate Forecasts?', *Proceedings of the National Academy of Sciences USA*, 4(99): 2487–2492.

Solomon, M. (2001) *Social Empiricism* (Cambridge, MA: MIT Press).

Stern, N. (2007) *The Economics of Climate Change: the Stern Review* (Cambridge: Cambridge University Press).

Wilholt, T. (forthcoming) 'Bias and Values in Scientific Research', *Studies in History and Philosophy of Science*.

Wimsatt, W. (2007) *Re-engineering Philosophy for Limited Beings: Piecewise Approximations to Reality* (Cambridge, MA: Harvard University Press).