Rationality in Inquiry

On the Revisability of Cognitive Standards

Jonas Nilsson
Rationality in Inquiry

On the Revisability of Cognitive Standards

Jonas Nilsson

Umeå 2000

Institutionen för filosofi och lingvistik
Umeå universitet
901 87 Umeå

Akademisk avhandling
som med vederbörligt tillstånd av
rektorsämbetet vid Umeå universitet
för avläggande av filosofie doktorsexamen
offentligen försvaras i hörsal F, humanisthuset
lördagen den 20 januari 2001, kl. 10.15
ABSTRACT
The topic of this study is to what extent standards of rational inquiry can be rationally criticized and revised. It is argued that it is rational to treat all such standards as open to criticism and revision.

Arguments to the effect that we are fallible with regard to all standards of rational inquiry are presented. Standards cannot be ultimately justified and with certainty established either as adequate or as inescapable presuppositions. Apel's attempt to give ultimate justifications of certain moral and logical rules is examined and criticized. Special attention is given to our fallibility with regard to logical inference rules. The idea that certain logical rules cannot be put into question because any critical argument presupposes them is criticized.

It has been claimed that there must be some basic standards which are such that they cannot be rationally evaluated and hence are rationally unrevisable. This is called "the unrevisability thesis". Related to this thesis is the normative policy according to which rationality requires that some standards be treated as unrevisable, the unrevisability policy. Two arguments that have been used to defend the unrevisability thesis and policy are examined and criticized. The conclusion is that we are not forced to accept either the thesis or the policy.

The negation of the unrevisability policy is the revisability policy, according to which it is rational to treat all standards as open to rational criticism and revision. Objections that have been directed against the revisability policy are discussed and criticized. According to the objections, the revisability policy leads to rationality relativism. These objections are refuted, and it is argued that it is, on the contrary, rational to adopt the revisability policy and treat all standards of rational inquiry as criticizable and revisable.

It is proposed that the rational change of standards should be viewed as a bootstrap process. General features of a bootstrap view of rational change of standards are presented, and it is argued that it is impossible to formulate a real theory of bootstrapping. Two models of standard change are presented and discussed: Laudan’s reticulated model of scientific rationality and Briskman’s bootstrap theory. It is claimed that in spite of defects and limitations, these models contribute to a richer understanding of bootstrapping.

The fallibility and revisability of standards of rational inquiry have consequences for how the normativity of rationality should be understood. The book ends with an account of how the rationality of cognitive actions is related to the idea of the adequacy of standards. A distinction between absolute and standard-relative rationality is made, and it is argued that what an inquiring agent rationally ought to do coincides with what it is standard-relatively rational for him to do. It is shown that this view of rationality of inquiry is nevertheless inconsistent with rationality relativism, and that it is compatible with an objectivistic view of rationality.

Keywords: Rationality, fallibilism, revisability, ultimate justification, relativism, bootstrap rationality, standards, cognitive goals, Laudan, Briskman, Worrall, Hauptli, Apel.
Rationality in Inquiry

On the Revisability of Cognitive Standards
Rationality in Inquiry

On the Revisability of Cognitive Standards

Jonas Nilsson

Umeå 2000
Till Mamma och Pappa
ABSTRACT

The topic of this study is to what extent standards of rational inquiry can be rationally criticized and revised. It is argued that it is rational to treat all such standards as open to criticism and revision.

Arguments to the effect that we are fallible with regard to all standards of rational inquiry are presented. Standards cannot be ultimately justified and with certainty established either as adequate or as inescapable presuppositions. Apel’s attempt to give ultimate justifications of certain moral and logical rules is examined and criticized. Special attention is given to our fallibility with regard to logical inference rules. The idea that certain logical rules cannot be put into question because any critical argument presupposes them is criticized.

It has been claimed that there must be some basic standards which are such that they cannot be rationally evaluated and hence are rationally unrevisable. This is called “the unrevisability thesis”. Related to this thesis is the normative policy according to which rationality requires that some standards be treated as unrevisable, the unrevisability policy. Two arguments that have been used to defend the unrevisability thesis and policy are examined and criticized. The conclusion is that we are not forced to accept either the thesis or the policy.

The negation of the unrevisability policy is the revisability policy, according to which it is rational to treat all standards as open to rational criticism and revision. Objections that have been directed against the revisability policy are discussed and criticized. According to the objections, the revisability policy leads to rationality relativism. These objections are refuted, and it is argued that it is, on the contrary, rational to adopt the revisability policy and treat all standards of rational inquiry as criticizable and revisable.

It is proposed that the rational change of standards should be viewed as a bootstrap process. General features of a bootstrap view of rational change of standards are presented, and it is argued that it is impossible to formulate a real theory of bootstrapping. Two models of standard change are presented and discussed: Laudan’s reticulated model of scientific rationality and Briskman’s bootstrap theory. It is claimed that in spite of defects and limitations, these models contribute to a richer understanding of bootstrapping.

The fallibility and revisability of standards of rational inquiry have consequences for how the normativity of rationality should be understood. The book ends with an account of how the rationality of cognitive actions is related to the idea of the adequacy of standards. A distinction between absolute and standard-relative rationality is made, and it is argued that what an inquiring agent rationally ought to do coincides with what it is standard-relatively rational for him to do. It is shown that this view of rationality of inquiry is nevertheless inconsistent with rationality relativism, and that it is compatible with an objectivistic view of rationality.

Keywords: Rationality, fallibilism, revisability, ultimate justification, relativism, bootstrap rationality, standards, cognitive goals, Laudan, Briskman, Worrall, Hauptli, Apel.
Acknowledgements

First I want to thank my supervisor Ingvar Johansson. He has been constantly supportive and generous towards me. The time and energy he has devoted to me and to my work, go far beyond what duty requires of a supervisor. His careful reading of everything I have written during these years, and his innumerable criticisms and suggestions, have been invaluable.

Two other Umeå philosophers have contributed in an especially important way, both to the writing of this dissertation and to my philosophical development in general. Gunnar Andersson has read and commented on what I have written, and I have learnt much from our discussions of philosophy of science and theories of rationality. Sten Lindström has also read and commented on my writings. Discussions with Sten have opened my eyes to many complications and difficulties that I would not have noticed otherwise, and they have greatly improved my understanding of the problems discussed in this work.

All of the chapters in this dissertation have been presented at the Friday seminar at the Philosophy department in Umeå. The discussions there have resulted in many improvements, and I want to thank all the participants at those seminars. In particular, I wish to thank Hildur Kalman, Per Nilsson, Peter Nilsson, Bertil Strömberg and Pär Sundström, who have all generously commented on different versions of the chapters presented here.

Our department secretary Ann-Katrin Minnhagen has helped me with many practical things, especially during the final stages of preparation for publication. For that I thank her.

I spent the autumn semester 1998 at the Philosophy department at York University in Toronto. I owe a special debt of gratitude to Professor Ian C. Jarvie, who invited me to York and supervised me while I was there. He generously provided me with an office of my own, and made many arrangements to make my stay there philosophically fruitful as well as pleasant. It was during that time I wrote the first drafts of most of the chapters of this dissertation, and our discussions together was a great help to me in coming to realize what it was I wanted to do and how I should do it.
I have had financial support for this work from the Philosophy department in Umeå, the Faculty of Humanities at Umeå University, and Riksbankens Jubileumsfond. I also received a grant from Kempefonden at an early stage of my graduate studies. While writing chapter 5, I received part-time financial support from Humanistiskt-Samhällsvetenskapliga forskningsrådet (HSFR) via the project Kunskap, kunskapsförändring och icke-monoton inferens (Knowledge, knowledge change and non-monotonic inference).

This book is dedicated to my parents Anita and Per-Ola.
Table of contents

1 Introduction
   1.1 Introduction 1
   1.2 Change and diversity in the history of philosophy 2
   1.3 Change and diversity in the history of science 3
   1.4 A normative concept of rationality 5
   1.5 Terminology: evaluative standards and methodological rules 8
   1.6 Conformance to standards: explicit and implicit standards 9
   1.7 Adequacy of standards and rationality of cognitive action 11
   1.8 Standard-relative and absolute rationality 12
   1.9 Rationality relativism 17
   1.10 The content and structure of this work 19

2 The fallibility of standards 21
   2.1 Introduction 21
   2.2 The Münchhausen trilemma and the impossibility of ultimate justifications 23
   2.3 Fallibility: the case of logic 28
   2.4 Does Apel avoid the Münchhausen trilemma? 31
   2.5 Apel’s attempt at ultimate justification 34
   2.6 A general epistemological objection against Apel’s justification procedure 36
   2.7 A counterexample to modus ponens 39
   2.8 Apel’s argument against fallibilism 44
   2.9 Conclusion 46

3 A critique of arguments for unrevisability 48
   3.1 Introduction 48
   3.2 Hauptli on the limits of evaluation 51
   3.3 The limits of justification argument 56
   3.4 Conclusion 61

4 A critique of arguments against revisability 63
   4.1 Introduction 63
   4.2 Worrall’s first objection: objectivity 65
   4.3 Worrall’s second objection: independence 68
   4.4 Unevaluability and relativism: turning the tables 72
   4.5 Conclusion 76
1
INTRODUCTION

1.1 Introduction
When we inquire into something, we try to effect some sort of cognitive change. We hope to learn something we did not know before, to understand something which puzzled us, to discover truths or get rid of falsehoods. In inquiries, we do a number of different things: we ask questions and raise problems, we argue, we formulate, test and revise theories and interpretations, we explore ideas and make decisions about what we should accept and what we ought to reject. In doing all these things we try, as best we can, to act rationally. We try to distinguish between mere change and rational change. By appealing to standards of rational inquiry we try to determine how we ought to behave in conducting our inquiries.

Most of the time our focus is not on rationality itself. We try to come to rational decisions about, for example, whether a certain theory offers an acceptable explanation of a given phenomenon. In doing this, we use or draw upon standards of rational inquiry, but rarely do we focus on these standards themselves and ask whether and how they can be evaluated and perhaps changed.

It is on questions of this latter sort this work will focus. The general question is whether we can be rational with regard to rationality itself. Can we be critical towards our standards of rational inquiry and evaluate such standards rationally? Or do we have to treat at least some of our standards as themselves being beyond rational evaluation, since we would always have to presuppose them in order to be able to evaluate anything at all?

Infallibilism — the view that we can know some things with absolute certainty — has become unfashionable in contemporary philosophy. Fallibilism, and the related idea that anything we presently believe or
accept may at some future point be revised, have come to seem attractive to many.

However, a number of philosophers have argued that when it comes to rationality, there must be absolute limits to what can be regarded as open to revision. Perhaps all general scientific and philosophical theories are fallible and revisable, and perhaps the same is true even for simple test reports and perceptual beliefs. But rationality, they maintain, must have fixed limits. It is arguments to this effect I shall examine in this work.

1.2 Change and diversity in the history of philosophy
Philosophers have studied, argued and disagreed about just about every subject matter one can think of. One thing philosophers have been concerned with has been how scientific inquiries should be conducted and what epistemological status the results of such inquiry have. But throughout the history of philosophy, philosophers have also had different ideas about how philosophical inquiry should itself be conducted if it is to be rational. It is glaringly obvious that the history of philosophy exhibits change and diversity, both in the standards and rules philosophers have explicitly proposed and in how they have conducted their inquiries. Think, for example, of the differences between the rationalist Descartes and the empiricist Hume; or between Kant’s critical philosophy and Hegel’s phenomenology. Think also of the diverse schools in 20th century philosophy: phenomenology, ordinary language philosophy, the philosophy of logical analysis, critical rationalism, various kinds of naturalism, and others. Interestingly, different and competing meta-methodological positions have developed within philosophy of science, offering conflicting ideas about how methodologies or methodological rules should be evaluated.

Indeed, precisely this change and diversity concerning ideas about how rational philosophical inquiry ought to be conducted has been taken to indicate that perhaps philosophy is not a rational kind of activity. Some have hoped, that when we turn to science it would be different. However, philosophers and historians of science have largely come to agree that, at least to some extent, change and diversity of standards and rules is to be found also in science.
1.3 Change and diversity in the history of science

Kuhn’s *The Structure of Scientific Revolutions*, when first published in 1962, inspired historians and philosophers of science to study scientific change not only of theories, but also at the level of methodological rules and standards. He was of course not the first to deal with the topic of changing standards of scientific inquiry. Indeed, discussions about methods and proposals of new methods for scientific inquiry have been common in the history of science. With their proposals for methodological reform scientists and philosophers such as Bacon, Galileo, Descartes, Newton and others saw themselves as contributing to the progress of science. But with Kuhn, the historical change of standards in science came to be seen as problematic. Kuhn held that the history of science should be understood as involving a sequence of paradigms, which include among other things their own paradigm-specific standards and rules of scientific rationality. His view was that rules and standards are internal to paradigms, and that they therefore change when one paradigm replaces another.\(^1\) Diversity of standards and rules was on the other hand an exception according to Kuhn. Such diversity is confined to periods of crisis, when no paradigm has yet been established as hegemonous normal science. Kuhn was interpreted as saying that changes of rules and standards (along with paradigm change) could not themselves be rational, and also that interparadigmatic debate and resolution of conflicts over standards and rules could not be rational.

Philosophers who wanted to do justice to the historical observations of Kuhn, but also wished to maintain the view that the history of science has been a largely rational affair, reacted in different ways to Kuhn’s historical arguments. Some thinkers, such as Lakatos, sought to develop general methodologies that were supposed to be neutral with respect to competing paradigms (or research programmes), and which could allow one to explain as rational the kinds of changes Kuhn described (see Lakatos 1974). Others, such as Laudan and Shapere, tried to explain how the standards and rules could themselves be objects of scientific debate

---

1 Although he insisted that adherents of different paradigms do share some standards and agree, for example, that theories should be accurate, simple, fruitful, and so on, such shared values can be both interpreted and weighted differently depending on what paradigm a scientist belongs to (see Kuhn 1970, 199).
and change rationally over time, and how they can be rationally debated so as to resolve conflicts over them.

It seems clear, anyway, that there have been changes in what standards and rules are used in scientific inquiry. Some examples are: the abandonment of teleological explanations; the introduction of the ideal of having mathematically formulated theories; the introduction of the idea of testing through active experimentation and systematic observation; the methodology of inductive generalization from observations was replaced by the method of hypothesis; single and double blind tests were introduced in testing the effects of clinical treatments and of pharmacological drugs; statistics was invented and so was the methodology of sampling and hypothesis evaluation that goes with it and which is still being developed; in physics the requirement that explanatory theories must be deterministic was abandoned.

I thus take it to be uncontroversial that standards and rules in science have to some extent changed. However, philosophers have disagreed about how such changes should be interpreted. The study of historical changes in standards and rules has motivated many philosophers to give accounts of how not only scientific theories, but also evaluative standards and methodological rules (as well as goals) may themselves be rationally changed. Whereas some have seen the changes as pervasive and have tried to formulate models in which all standards and rules are seen as open, in principle, to revision (see Laudan 1984 and 1996; Shapere 1984), others have instead argued that such changes as have been are peripheral and that they can be explained as rational (or irrational) on the basis of some core set of standards and rules that has remained constant (Worrall 1988; Newton-Smith 1981).

It is not my ambition to evaluate historical claims about what changes there have been and how extensive they are, or to settle the issue of how particular changes are best explained. Instead I shall discuss more abstract questions about rationality in inquiry, questions that concern what is in principle open to change and what is not in our understanding of what

---

2 Proponents of the method of hypothesis maintained that general statements did not have to be arrived at by generalization from observations: it was permissible to propose hypotheses first and to confirm them afterwards, by testing their empirical consequences. This method is nowadays often called the "hypothetico-deductive method". For extensive treatments of this case, see especially Laudan 1981, and also Gower 1997, in particular ch. 6.
such rationality is. Can one for example prove with certainty that some particular standards or rules are adequate or that they are indispensable for rationality? Or are we fallible with regard to all standards and rules? Are there reasons for thinking that there are standards that are forever beyond the possibility of criticism and change? Or are there reasons for thinking that they are all open to change? Is it a rational policy to set definite limits to rationality and treat some standards as uncriticizable and unrevisable? Or is it perhaps a more rational policy to treat them all as open to criticism and to rational revision?

It is questions at this very general level I shall discuss in this work. They are only indirectly related, I think, to more substantive questions about whether particular changes in the history of science or philosophy were rational or not, and to questions about which particular standards and rules are actually adequate and to be recommended for various kinds of inquiry.

1.4 A normative concept of rationality
Rationality is here used as a normative concept. It is used in evaluating actions, and to issue normative recommendations about how agents should act. To say that an agent is acting rationally is to appraise his action positively, to approve of it in some way.

I shall here limit myself to what may be called cognitive rationality, and speak of what inquiring agents rationally ought to do insofar as they are that: agents engaged in some form of inquiry. Thus I leave the possibility open that some other form of normativity, of a moral or prudential nature, must also be taken into consideration in order to determine what an agent, all things considered, ought to do. When I write about what an inquiring agent rationally ought to do, I intend this as an abstraction that only takes into consideration things that are so to speak internal to the inquiry in question. It is possible that some particular action is what one rationally ought to do (on the basis of the goals, standards, information and arguments within a particular inquiry), but that there could be for example moral reasons for acting otherwise (e.g., refusing to perform certain kinds of experiment or even to accept a certain theory). I shall not deal with these complications here, though.
Agents who try to act rationally, do so by trying to conform to standards and rules they take to be appropriate for guiding particular kinds of cognitive actions. Similarly, in judgments of rationality and irrationality appeal is made to normative standards and rules that the actions in question conform to or violate. But rationality cannot be identified with conformance to some particular standards and rules. That an action is rational does not mean that it is such that it conforms to such and such standards. To try to analyze the meaning of "rational" or "rationality" in that way (for example by specifying the relevant standards conformance to which is necessary and sufficient for an action being rational) will automatically fail. The reason is that it would amount to giving a descriptive analysis — "conforms to such and such standards" — of a normative concept. The normative element of rationality would thus be left out, the element that consists in an attitude of positive appraisal of the action as what ought to be done.

Thus if one proposed that "rational" means something like "in accordance with standards S" — where S might be for example Descartes’ rules for the direction of the mind, or Bayesian methodology, or Popper’s falsificationism — one might always intelligibly ask: "But is it really rational to act in accordance with S?" This is in effect an instance of the "open question argument" that G.E. Moore used in Principia Ethica (1903) to argue against naturalistic analyses of "good".3

Now, it might be correct (perhaps even true) that a cognitive action is rational if and only if it conforms to some theory of rationality such as the ones mentioned above, but that would not amount to a definition of, or an analysis of the meaning of, "rationality".

It is impossible to give a definition which states necessary and sufficient conditions for the application of the concept of rationality. I shall certainly not attempt such a definition myself. Instead I shall view methodological theories and particular standards and rules not as attempts to define "rational" but as claims about how one rationally ought to act. They make substantive claims about what is rational, claims that may or may not be correct but which can be intelligibly questioned.

---

3 Gibbard relates Moore’s argument explicitly to rationality. See Gibbard 1990, 11. See also Darwall, Gibbard and Railton 1992 for a detailed discussion of the open question argument and the alleged naturalistic fallacy.
It is of course possible to use the term "rational" as descriptive, and stipulate that the term means, for example "maximizes expected utility" or "conforms to falsificationist methodology" or "dispassionate and calculating". This would be to turn the ordinary, normative concept of rationality into a descriptive concept by means of a stipulative definition. It is always possible to do this, but one should bear in mind that in doing so one changes the subject and is discussing a quite different concept. When the term is used descriptively in this way it is severed from the question of how one ought to act. I believe that most criticisms and rejections of rationality are best understood as employing a descriptive and thus non-normative concept like this. But it should be noted that the normative rationality question can always be raised concerning any descriptive characterization of "rational": Is it always irrational to act in a way which is known not to maximize one’s expected utility? Is it really always a good thing to be dispassionate and calculating when trying to figure out what to do — might that not be too constraining and irrational in some cases? These are normative questions and one does not settle them by making a stipulative definition of "rational".

Another possible type of case where one might say that one ought not to be rational — for example that one ought to believe something although it is not rational to believe that — is if what cognitive rationality requires (what one as an inquiring agent rationally ought to do) is taken to be in conflict with what morality (or perhaps one’s well-being and prudential rationality) requires. I do not want to rule out the possibility that there might be conflicts between what is cognitively rational and what is rational all things considered, or that there might be conflicts between what is rational and what is morally permitted. But I shall not pursue the question of how rationality in inquiry is related to prudential rationality and other kinds of normative considerations.

---

4 Nozick distinguishes between "p being the rational thing to believe" (cognitive rationality) and "believing p being the rational thing to do" (prudential rationality) and tries to integrate the two (Nozick 1993, 70).
1.5 Terminology: evaluative standards and methodological rules

Inquiring agents use various normative things to guide their various cognitive actions. I shall refer to all of these by the generic term "standards of rational inquiry". These things, the standards, have been called a lot of things: "methods", "principles", "criteria", "rules", "norms", "desiderata", "standards", etc. They are supposed to guide certain aspects of inquiry, and range from logical inference rules and general methodological injunctions and standards of theory evaluation, to principles or standards specific to particular disciplines or approaches, and on down to detailed prescriptions for correct handling of laboratory equipment.

I shall use the terms "evaluative standards" and "methodological rules" to refer to two subclasses of this broad class of standards of rational inquiry.

By "evaluative standards" I mean standards that are used for evaluating things that are not themselves actions, such as theories or hypotheses. Standards of theory choice are of this type, such as "a successor theory must include all the corroborated content of its predecessor", "prefer simple hypotheses to ones that are less simple", "prefer interpretations that maximize the coherence of the work being interpreted", "accept only explanations where the existence of the causes appealed to can be independently verified", and so on. But within "evaluative standards" I also include standards that are applied to standards and rules (meta-standards), such as "a set of standards must be consistent", "among alternative standards, choose that one which you believe to be the most efficient means to your cognitive goals", or "standards must be consistently self-applicable". These evaluative standards are used to evaluate something which is not itself an action.

By "methodological rules" on the other hand, I shall mean something like norms or rules that tell agents to act in particular ways, and which are not used for evaluating things other than actions. They tell you to avoid certain kinds of actions ("do not try to save theories from refutation") and recommend other kinds of actions ("test your theories as severely as you can"). Heuristic rules belong in this category.

I am aware that the distinction made here is not very sharp. Larry Briskman, from whom I borrow the distinction (see Briskman 1977),
writes that every evaluative standard can be reformulated as a methodological rule, and vice versa. An example might be the standard "prefer independently testable hypotheses to ones that are not", and the corresponding methodological rule "avoid the introduction of hypotheses that are not independently testable". I am not sure if this is true of every evaluative standard and methodological rule. Furthermore, one could say that methodological rules are in a sense evaluative standards: such rules can be used to evaluate actions that conform to or violate the rules. But I hope the distinction is nevertheless clear enough for the purposes I put it to in this work.

Finally, I shall sometimes discuss logical inference rules (rules of deductive logic). I shall treat them as a separate subclass of the genus "standards of rational inquiry".5

The discussion in the chapters to follow will largely be focused on evaluative standards, but in many cases what I say will apply equally to evaluative standards, methodological rules and logical inference rules, and then I will just use the term "standards" to refer to the genus standards of rational inquiry.

1.6 Conformance to standards: explicit and implicit standards
Rationality, as used here, is something we ascribe to individual actions performed by agents — not to, for example, theories, actions taken in abstraction, or agents.6

A rational action presupposes, one could say, a will to be rational. According to the conception of rationality used here, to say that an agent is acting rationally or irrationally implies that the agent is guided, in his action, by normative considerations about how this kind of action rationally ought to be performed. It implies, that is, that the agent is

5 Logical inference rules can be viewed as evaluative standards that are used for evaluating arguments. Alternatively, they can be viewed as methodological rules that permit you to make certain inferences. How they are classified is not important for my purposes, though.

6 "That is rational" can be analyzed in different ways. One can think of it as a judgment that attributes a property, rationality, to an action. On the other hand, an expressivist like Gibbard would deny that to say of an action that it is rational is to ascribe a property to it. According to Gibbard, what one does is instead to express one's acceptance of norms that permit the action (Gibbard 1990, 6-8).
trying to conform to standards. I shall say that an agent who tries to conform to a standard takes that standard to be adequate.\(^7\)

I am here using a distinction between conformance to and accordance with standards and rules. I borrow the terms from Pollock. In his theory of epistemic justification he says the following concerning epistemic norms:

To say that you act in accordance with a norm is just to say that your behavior does not violate the norm. This is compatible with your doing it for some reason unrelated to the norm. To say that you act in conformance with the norm is to say not only that you act in accordance with the norm but also that your behavior is guided by the norm. Justification requires conformance — not just accordance. (Pollock 1986, 168)\(^8\)

In performing a certain action, such as deciding that a theory \(T_1\) is better than \(T_2\), an agent might be guided by a wholly inadequate standard, or he may be misapplying a standard. At the same time it might be that by adequate standards \(T_1\) should indeed be judged as better than \(T_2\). The result might be that he so to speak ”gets it right” so that his action accords with what an adequate standard recommends. But he would not be conforming to that standard, and his action might be irrational although it accords with the adequate standard. If you will, it would be an instance of doing the right thing for the wrong reason.

I want to emphasize that I am not presupposing an ”intellectualist” model of action here. That an agent in acting tries to conform to a standard or rule does not mean that he makes an explicit appeal to an explicitly formulated standard. The standards involved may be, and often are, implicit. Agents can be guided, in Pollock’s sense, by a norm or standard without thinking of a formulation of the standard. Actions can therefore conform to standards even if the latter are implicit. Agents who

\(^7\) The concept of adequacy is discussed in section 1.7.
\(^8\) This can be compared to C.D. Broad’s threefold distinction between ”action in accordance with, action on, and action for a principle” (Broad 1930, 119). Acting on a principle in Broad’s sense would be conforming to it in Pollock’s sense. So would acting for a principle be. But in Broad’s terminology, only actions ”done on a principle which is a categorical imperative” are actions for a principle. Pollock’s ”conformance to” applies to both hypothetical and categorical imperatives.
use deductive reasoning, for example, are often unable to formulate the logical rules they are using. Still that reasoning is rule-governed, and although they have not formulated the rules that guide their action, they can often recognize, for example, when they make mistakes in their reasoning. Here, I shall not try to describe how implicit standards can guide actions, but merely assume that they can do so.9

I shall, however, focus on explicitly formulated standards. For obvious reasons, it is easier for me to discuss these. More importantly, in order to criticize standards and show them to be problematic, and to compare standards in cases of change or conflict they must somehow first be articulated and made into something that can be the object of a critical discussion.

1.7 Adequacy of standards and rationality of cognitive action
So rationality is a property of (cognitive) actions performed by (inquiring) agents. Their actions are guided (whether successfully or not) by normative standards. According to my terminology, standards and rules are not themselves rational or irrational. They belong, like theories, to the wrong category for the concept of rationality to apply to them. I shall instead say that standards and rules are adequate or inadequate. Alternative terms might have been "good", "correct", "valid", or "appropriate".

Adequate are what standards should be, then, and adequacy is what is claimed for the standards that agents use to normatively guide their cognitive behaviour. In criticism one can criticize an action as irrational (or not fully rational), by arguing that it violates standards taken to be adequate by the critic. A standard, in turn, can be criticized for being inadequate. It can be rational or irrational to accept or use a particular standard, but the standard itself is neither rational nor irrational.

To say that a standard is adequate or inadequate is itself to make a normative claim. What is it, then, for a standard to be adequate? What does the goodness of standards consist in?

The majority view among philosophers is that the adequacy of evaluative standards and methodological rules is a matter of the extent to

---

9 See, for example, Pollock's discussion of how implicit standards guide cognitive behaviour (Pollock 1986, 127-32).
which they are suitable means for achieving the goals of inquiry. According to this view, standards are instruments and their adequacy consists in their being efficient means in cognitive goal-pursuit. This provides a rationale for adopting and using standards — we use them in inquiry because we believe that this helps us achieve our cognitive goals. It also gives some content to the idea that one standard can be more adequate, or better, than another one: it is more adequate if it better enables us to achieve the goals of inquiry.

Epistemologists and philosophers of science who embrace this instrumentalist conception of standards usually have definite opinions about the goals to which the standards are means. It is such a broadly instrumentalist idea that epistemologist Susan Haack assumes when she writes that "truth-indicative is what criteria of justification need to be to be good" (Haack 1993, 203), and it is clearly found in reliabilist accounts of justification such as that of Alvin Goldman. According to Goldman a belief is justified just in case it is produced by a reliable process or method of belief acquisition. That a process or method is reliable means that it leads to a high frequency of true beliefs (see e.g. Goldman 1986). Thus for both Haack and Goldman, whether a standard is adequate or not depends on whether it is a good means for achieving the goal of having true beliefs.

Myself I believe that this instrumentalist view, that takes the adequacy of standards to be a matter of goal-directedness, is essentially correct. Thus, in my opinion a standard is adequate to the extent that it is an appropriate means to cognitive goals. I shall, however, later on add some reservations and qualifications to this instrumentalist claim, especially in chapters 5 and 6. Above all, different agents might decide to adopt different goals, and a theory of rationality in inquiry should also say something about whether these goals can themselves be rationally evaluated and how this is to be done.

1.8 Standard-relative and absolute rationality
Inquiring agents try to conform to normative standards and rules as they conduct their inquiries. Similarly, when one judges that an agent is acting rationally or irrationally one does so by appealing to normative standards and rules and seeing whether his action conforms to it or not. I have
already said that for an action to be rational the action must conform to a standard and not merely accord with it. But the standards to which an agent is trying to conform in, for example, comparing and choosing between two theories, may not be adequate. He may accept standards that are inadequate, without his knowing so. Someone observing the agent may judge that his action is guided by inadequate standards, and that it violates standards that are adequate.

An important theme in this work is that it is always possible that some standard of rational inquiry that we have taken to be adequate is in fact inadequate — we are fallible with regard to the adequacy of our standards. Two considerations apparently point in different directions when it comes to trying to understand the normativity of rationality.

On the one hand, there is an intimate connection between the rationality of actions and the adequacy of standards. An agent who is guided by a standard takes that standard to be, in some sense, adequate. It is taken to be a standard which not only is accepted by him (or in his culture or in his discipline or in his research team) but which is also believed to be adequate: it is a good standard and one ought to conform to it whether or not one in fact does so. When someone’s actions are evaluated in terms of rationality and irrationality, the question is normally seen as one of whether or not the action conforms to or violates the adequate standards that are relevant for the case in question. And that is also what the acting agent, who is being evaluated, is himself aspiring to: to make his action one that conforms to adequate standards. This suggests that rationality is a matter of acting in conformance with adequate standards.

On the other hand, rationality is supposed to be capable of guiding agents in their actions. An agent tries to figure out what the rational thing to do in a particular situation is (for example, whether he should accept or reject a theory, or if one theory should be preferred as better than another one). To be properly action-guiding, rationality, what an agent rationally ought to do, must be something that is attainable for the agent in the situation he is in. For example, an agent rarely has access to all the information that is relevant for a decision he is about to make. One can of course ask what would be the rational thing to do if one had all the information. But one cannot act on information that one does not have. The agent (and any external evaluator of his action) is still left with the question of what decision it is rational to make in view of the information
that he does have. To say he ought to act in the light of all the information that bears on the decision, is to propose a standard that he cannot use as a guide for his action. Rationality, as has often been emphasized, is a matter of doing the best one can with the resources at one's disposal — and these resources may be more or less limited. Furthermore, for an action to be rational it is not enough that it happens to accord with standards. The agent, in acting, has to *conform* to the standards in question — the standards must be ones that the agent accepts and tries to guide his action by. But what happens when the agent in question is mistaken about what standards are adequate — if he accepts and tries to conform to standards that are in fact, without his knowing so, not adequate? Given that he accepts the standards he does, what ought he to do?

Say that the agent accepts the inadequate set of standards $S_1$ and that conforming to $S_1$ would be to perform the action $A_1$, whereas according to the set of standards $S_2$ that are in fact adequate another action $A_2$ should be performed. How should this agent act? It is plausible to say that the rational thing to do is to perform $A_2$ which is what the adequate standards recommend. At the same time, if the agent imagined here were to do that, although he would act in accordance with adequate standards, he would also violate his own standards $S_1$, the standards he himself takes to be adequate. Would that really be rational of him? I suggest that that would instead be a paradigmatic case of irrationality.

The actions of an agent can be described in different ways. One can describe an action as according with a certain set of standards. Whether these standards are adequate or not is another matter. One can also describe an action as being in conformance to a set of standards, which implies that the agent in question accepts these standards and that his action is guided by them. Again, the question of whether the set of standards is adequate may be left open. Whatever standards an agent accepts, his actions can be described as conforming to or violating his own standards. This is just stating facts about his actions. How they should be evaluated, whether the agent is acting rationally, is a different question. When the standards to which the agent is conforming are indeed adequate, the matter is straightforward. He is acting rationally. He is doing what he rationally ought to, which is here what adequate standards require. The move from describing the action as conforming to a set of

---

10 See, for example, Rescher 1988, 1-2.
standards, to the normative judgment that the agent acts rationally, here has to do with the normative judgment that the standards involved are adequate.

Let us say that an action which conforms to standards that are adequate is *absolutely rational*. This concept of absolute rationality is clearly a normative concept. Next let us say that an action which conforms to the standards accepted by the acting agent is *standard-relatively rational*. Note that here no assumption is made about whether the agent’s standards are adequate or not. Should this concept of standard-relative rationality be viewed as a normative concept? Does it have anything to do with how the agent rationally ought to act, or is it just a descriptive concept?

Note first that standard-relative rationality is a necessary condition for absolute rationality. Conformance to a standard implies that the agent who is conforming in some sense accepts that standard as adequate. So every case of absolute rationality is a case of standard-relative rationality, but not vice versa.

But when an agent acts in conformance to inadequate standards, and is *merely* standard-relatively rational, it is perhaps not so clear that this fact of conformance is normatively relevant. Indeed, actions that go against what adequate standards of rational inquiry require are not something one would want to recommend — as external observers we would regard such an agent as making a mistake. I shall nevertheless argue that standard-relative rationality *is* a normative concept, and that it has to do with what agents rationally ought to do, even in cases where the standards they accept happen to be inadequate.

The standards an agent accepts as adequate may be seen as resources at his disposal, just like information. Some of the information he has may be mistaken, and some of his standards may be inadequate. Having to act on mistaken information and inadequate standards is unfortunate. In either case, it may be a result of past irrationality, but it need not be. Still, an agent has to decide how he should act, and given the information he has and the standards he does believe to be adequate, what he rationally ought to do may amount to acting in conformance to inadequate standards. In some cases at least, a standard-relatively rational action will be what the agent rationally ought to do, even though that action is not absolutely rational.
Assume for example that a scientist is working in a field where two different theories, $T_1$ and $T_2$ are competing. From $T_1$ (which could be a wave theory of light) a new and surprising prediction is inferred: it predicts that when a disc is placed at a certain distance from a source of light a bright spot will be visible in the middle of the shadow cast on a screen by the disc. The surprising prediction turns out to be correct. $T_2$ (which could be a particle theory of light) did not predict this, so it is now in trouble. However $T_2$ is then revised into $T_2'$ so that the new particle theory now correctly predicts the bright spot on the screen. Otherwise, both $T_1$ and $T_2'$ can explain the facts known so far. Assume further that our scientist accepts something like what has sometimes been called the "rule of predesignation", which says that a theory is better supported by correctly predicting phenomena of a kind that were unknown and surprising when the theory was formulated, than by merely predicting already known kinds of phenomena. Something like this evaluative standard has been advocated by many scientists and philosophers, but there are others who reject it (see Laudan 1984 for references and discussion; also Worrall 1989). Our scientist, who accepts this standard, realizes that according to this standard $T_1$ is better supported than $T_2'$ and should be preferred. So he prefers $T_1$ to $T_2'$. As far as this standard goes, he is standard-relatively rational. But now suppose, hypothetically and for the sake of the example, that this standard, or the rule of predesignation, is in fact not adequate. Suppose that by correct evaluative standards, it is irrelevant whether phenomena predicted by a theory are of a kind unknown at the time of its formulation or not, and that by adequate standards the rational thing to do would be to judge that the two theories have the same degree of support and that neither one should be preferred to the other. But our scientist does not know this. Let us assume that he has not been presented with any arguments for thinking that this standard is problematic. He accepts his standard as adequate and believes that it discriminates clearly in favour of $T_1$. What ought he to do in this situation? What is the rational thing to do? I suggest that in this case, he would be acting irrationally if he failed to conform to the inadequate standard. He does not know that his standard is inadequate and has no particular reasons to suspect that it is, so he can not take account of this as he figures out what it is rational to do. Given what he believes about how theories should be evaluated, the best thing he can do in this
situation is to try to conform to the evaluative standards he thinks are adequate, and not to do so would be to fail to perform the action he himself believes is the best one in the situation as he sees it.

In a case like this, to perform the standard-relatively rational action, which is not absolutely rational, is what the agent actually ought to do. To deviate from this would be irrational, even if it would have been better if he had known what the adequate standards are and applied these instead. In the situation he is in, he does not know this, and being standard-relatively rational is the best he can do, and what he rationally ought to.

It is in the agent's interest that his standards be adequate and that, if he presently has inadequate standards, to detect this and revise his standards (just as it is in his interest to discover new standards that allow him to inquire in new and better ways than before).

Still, in trying to criticize his own standards to detect possible inadequacies, and in judging about whether one set of standards is more adequate than another one, an agent must try to be standard-relatively rational. This places some restrictions on the ways in which standards of rational inquiry can be rationally criticized and changed, which will be discussed in subsequent chapters.

In the concluding chapter 6 I shall discuss in more detail the relation between the adequacy of standards and what agents rationally ought to do.

1.9 Rationality relativism
In this work I will argue that what it is rational for inquiring agents to do depends in an important way on what standards the agents in question actually accept. This means that I hold that rationality in inquiry is in a sense relative to the sets of standards of rational inquiry accepted by particular agents at particular times.

However, I want to make it clear that the thesis that rationality is relative in this way, does not amount to relativism. Indeed, although I shall not argue directly against relativism in this work, I regard it as a condition of adequacy on an account of rationality that it does not have relativism as a consequence.

---

11 For such arguments, see e.g. Siegel 1987, which has the title Relativism Refuted.
There are many different kinds of relativism on offer, depending on what is claimed to be relative. In discussions in epistemology and in the philosophy of science, the relevant kind of relativism is often taken to be one which relativizes either rationality or truth, or both of these. Here I shall not discuss relativism of truth. Instead, my worry concerns what I will call rationality relativism.

Rationality relativism would say not only that different agents can accept different sets of standards of rational inquiry or that one agent can accept different standards at different times, or that what is rational for an agent depends on what his set of standards are. The kind of rationality relativism I am interested in avoiding adds something more to this. It also consists of some of the following claims: that standards themselves cannot be adequate or inadequate; that there is no point in claiming that one standard (or set of standards) is better than another one; that standards (or a set of standards) cannot be rationally criticized; that agents who accept different standards do not really disagree about what is rational since in reality they have different rationalities.

An important aspect of this rationality relativism is the claim that it is impossible to be rational with regard to one’s standards of rationality. If that were true, the topics discussed in this work would not be real: how can standards of rational inquiry be rationally criticized? how can standards change rationally? and how can rational discussion lead to resolutions of conflicts about rationality? If rationality relativism were true then questions such as these would be, at best, misguided.

In chapter 6 I will give a more precise characterization of this rationality relativism, and discuss how the different claims I will make about rationality in inquiry are related to that relativism. But it will also be referred to in many arguments that I discuss in the following chapters (this concerns both arguments by others which I examine and arguments that I myself make). In these arguments relativism is that unacceptable consequence that different positions or theses are argued to lead to — it is the absurdity which positions are criticized for having as a logical consequence.

---

12 For an overview of different kinds of relativism, see Harré and Krausz 1996.
1.10 The content and structure of this work

In chapter 2 I argue that we are fallible with regard to all standards of rational inquiry. By this I mean that it is not possible to give ultimate justifications for claims that standards are adequate. An ultimate justification would enable us to establish with certainty that a standard is adequate, and rule out any possibility, even in principle, of being mistaken about it. I argue that, quite generally, any attempt to give an ultimate justification for a standard leads to a form of what Hans Albert has called "Münchhausen's trilemma". Someone trying to give such a justification ends up having to choose between three alternatives: an infinite regress of justifications; a viciously circular justification; or a justification claimed to end with something purportedly self-justifying. Neither of these alternatives can lead to the kind of justification that is sought. This applies to any attempt to give ultimate justifications of standards, and if my objections are correct then we are fallible with regard to all standards. However, the chapter also contains a detailed examination and criticism of Karl-Otto Apel's attempt to provide ultimate justifications of basic standards of rationality and ethical rules by means of "transcendental-pragmatic reflection". I argue that Apel does indeed fail to give ultimate justifications. I also criticize the idea that some logical inference rules must be thought of as inescapable and unreviewable presuppositions of rationality because they would be presupposed in any attempt to criticize them.

In chapters 3 and 4 I examine arguments that have been claimed to show that, although standards cannot be ultimately justified, we must still regard rationality as being necessarily limited in certain ways — that there must be standards of rationality that cannot themselves be rationally scrutinized. Chapter 3 deals with arguments to the effect that some standards are such that we need them in order to be able to argue and rationally evaluate anything at all, whereas these standards themselves cannot be subjected to critical evaluation. A consequence of this would be that it is also impossible to rationally revise these standards. I try to show that these arguments are mistaken, and that we are not forced to believe that rationality is thus limited.

In chapter 4 a related argument is examined. That argument is to the effect that anyone who adopts a policy of treating all standards as being open to rational criticism and revision, is forced into an untenable
relativist position. I argue that this policy does not have this relativistic consequence. As far as avoiding relativism goes it is on the contrary preferable not to treat any standards as uncriticizable and unrevisable.

The upshot of the mainly negative argumentation in these chapters is a view of rationality in inquiry as something which is itself open to critical discussion and possibly to rational change. The two chapters that follow are devoted to discussions of some consequences of this view of standards of rational inquiry as being fallible, criticizable and revisable.

In chapter 5 two main questions are addressed. On the assumption that no standard is to be treated as immune to revision, what does the process of evaluating and improving standards look like? And is it possible to formulate a general theory about this process? I propose that the process of evaluating and improving standards is best seen as a bootstrap process. However, a result of viewing it as a bootstrap process is that it is not something one can have an informative and interesting theory about. If it is a matter of bootstrapping, then this process is in principle open-ended and cannot be legislated in advance.

In the final chapter, the focus is on the problem of how the possibility of change and diversity of standards can be reconciled with a normative and non-relativistic conception of rationality. The distinction between standard-relative and absolute rationality is further elaborated. I introduce the concept of an agent's subjective rationality duty and argue that what this duty requires of an agent — what he rationally ought to do — is that he be standard-relatively rational by conforming to the standards he himself accepts. Thus, what an agent rationally ought to do is relative to the set of standards he accepts. At the same time, an agent aspires to have standards that are adequate, and it is in his interest not only to do his subjective rationality duty but also to try to detect and revise possible inadequacies among his standards. In trying to criticize and improve standards, it is helpful to have a specific conception of what the adequacy of standards consists in. However, I also argue that any such conception of adequacy is itself criticizable and revisable. Even at this level, rationality is something we can and ought to treat as open to criticism and revision.
THE FALLIBILITY OF STANDARDS

2.1 Introduction
In this chapter I argue that we are fallible with regard to standards of rational inquiry. I argue that standards cannot be ultimately justified. By this I mean that they cannot be justified in a way which rules out the possibility that they are inadequate. In this sense we are fallible in relation to standards. We cannot know with absolute certainty that a standard is adequate.

It is not the standards themselves that are fallible. It is agents who are fallible with regard to standards, who may be mistaken about whether standards are adequate or not. More precisely, inquiring agents cannot provide ultimate justifications of standards — cannot rule out the possibility of being mistaken about their adequacy — and thus cannot know with certainty that standards are adequate.

In discussing whether fallibilism extends to logic, Susan Haack distinguishes between "proposition fallibilism" which concerns the possible falsity of propositions, and "agent fallibilism" which is the "epistemological thesis that we are liable to hold false beliefs" (Haack 1978, 234). Her point is that even if logical laws are necessary, and thus cannot possibly be false, this is compatible with agent fallibilism: we may be mistaken about which the necessarily true laws are. As she puts it, "proposition infallibilism doesn't entail agent infallibilism" (Haack 1978, 234).

It is such an agent fallibilism I propose concerning standards, although I formulate it more generally in terms of the adequacy of standards rather than the truth of logical laws. It is we (all inquiring agents) who may be mistaken about whether or not standards are adequate. This does not mean that all standards are possibly inadequate. Our fallibility is compatible with there being standards that are necessarily adequate (or necessarily
inadequate). If we assume, for example, that all valid logical inference rules are adequate and that if a rule is valid then it is necessarily valid, then valid logical inference rules would be examples of standards that are necessarily adequate. That there are such standards I do not deny. What the fallibilism I propose denies, is that we can know with certainty that a standard is adequate.

Of course, if a standard is necessarily adequate, and if we think it is adequate, then it is not possible that we are mistaken (just as in believing a necessary truth it is not possible to be mistaken). But we cannot know with certainty that a standard is adequate, nor that we are right in accepting it as adequate.

Some attempts at ultimate justifications do not aim to show directly that standards are adequate, but rather that they are inescapable presuppositions of rationality. I shall consider and criticize such attempts, too. Hence in this chapter I use "fallible" in a broad sense, so that it means that standards cannot with certainty be justified as either adequate or inescapable.

The fallibilist view I try to defend concerning standards is largely in agreement with, and inspired by, the general fallibilism that is part of the critical rationalism of philosophers such as Hans Albert and William W. Bartley. However, I am only concerned with arguing that we are fallible with regard to standards. I shall not discuss whether this fallibilism can be extended to other things as well.

I begin by presenting a general objection against the idea that standards can be ultimately justified, and go on to discuss some reasons that have been offered for thinking that we are fallible with regard to logical inference rules. Then I consider Karl-Otto Apel's attempt to give ultimate justifications of certain principles of rationality and ethics, and argue that they fail. At the end of the chapter I discuss in more detail the idea that logical inference rules must be seen as inescapable presuppositions of rationality — an idea that seems to be crucial in arguments for infallibilism.

---

1 One might of course connect the concepts of inescapability and adequacy. One could hold, for example, that if a standard is an inescapable presupposition of rationality, then it is constitutive of rationality and therefore necessarily adequate. Furthermore, an ultimate justification which shows that a standard is inescapable would also show the standard to be un revisable, in the sense used in chapter 3.
2.2 The Münchhausen trilemma and the impossibility of ultimate justifications

The purpose of this section is to argue that anyone who attempts to give ultimate justifications of standards faces a general difficulty which cannot be overcome. He would be trapped in what Hans Albert calls "the Münchhausen trilemma".

In *Treatise on Critical Reason* (1985), Hans Albert argues simultaneously against the possibility of ultimate justifications and against the principle of sufficient justification. That principle he expresses as follows: "always seek an adequate foundation — a sufficient justification — for all your convictions." (Albert 1985, 14) An alternative formulation is the following: "It is rational to accept something if, and only if, it is sufficiently justified." On Albert’s view, this principle is connected to a strive for certain knowledge, which rests on secure foundations. That is, Albert seems to suggest that justificationism requires the availability of ultimate justifications, which are to provide an Archimedean point for knowledge. I do not think that this is the case — one might accept the principle of sufficient justification, without requiring that justified knowledge must be certain. In this chapter, I will therefore discuss only those parts of Albert’s argumentation that concern the possibility of ultimate justifications.

An ultimate justification is supposed to show that the truth or adequacy of a particular item of knowledge is absolutely certain, so that we are infallible with regard to it. It must rule out the possibility, in principle, of being mistaken about the object of the justification.

But this cannot be done, according to Albert. Any attempt to achieve an ultimate justification, he claims, leads to what he calls the "Münchhausen trilemma". The trilemma is a situation in which one is forced to choose between three alternatives, neither of which is adequate for achieving an ultimate justification (Albert 1985, 18).

I shall reformulate the trilemma slightly, so as to deal explicitly with attempts to give ultimate justifications of standards of rational inquiry. For the sake of brevity I shall speak mainly of "standards". My claim, which builds on Albert’s argument about knowledge in general, is that
anyone who attempts such a justification is faced with a choice between the following three alternatives:²

1. One might try to prove, or in some other way give justifying reasons for the standard. But if the justification is to be ultimate (and remove the possibility of error once and for all), then that which is invoked as premises in the proof or as some other kind of justifying reason — as well as the claim that the mode of inference used (deductive or whatever it may be) is adequate — must also be justified. If one tries to do this by providing further justifications, then everything appealed to in giving these new justifications must itself be justified. And so on. This leads to an infinite regress.

2. One might escape the regress by resorting to a circular justification — for example by using an unjustified standard A to justify standard B, and then use B to justify A. But although the result of this kind of operation might be coherent, it is inadequate if the justification is to be ultimate.

3. Finally, one might try to end the chain of justifications by holding that the standard is such that it simply does not require to be justified by anything else, that it is self-evident or self-justifying, and that its correctness can be grasped in an immediate way, without its having to be further justified through anything else.

An example of an infinite regress, would be if I tried to justify a logical inference rule by giving a meta-logical proof of it (by showing that there can be no counterexamples to the rule), and then tried to justify this proof by giving a meta-meta-logical proof for it, and so on.

As an example of a circular justification, one can perhaps take Descartes’ attempt, in the Meditations, to justify a standard which says that what I perceive clearly and distinctly is guaranteed to be true. To rule out the possibility of being mistaken about that which is clearly and distinctly perceived, he sets out to prove the existence of a benevolent God. The existence of God justifies the standard, since his benevolence is incompatible with my being thus mistaken. However, in order to prove

² The trilemma is a version of a well known epistemological argument, sometimes referred to as the "regress argument". It has traditionally been used to argue for foundationalism, which amounts to adopting some version of the third alternative. Susan Haack has recently reformulated and criticized this argument, which she calls the "No Tolerable Alternatives Argument" (Haack 1993, 21-25). The reason Albert calls this a "trilemma" is of course that he regards the third alternative as untenable, too. Also, see the limits of justification argument discussed in section 3.3.
that God exists, Descartes has to rely on the very standard that God’s benevolence is supposed to justify. God justifies the standard, the standard justifies the proof that God exists.\(^3\)

Since attempts at ultimate justifications usually consist in choosing the third alternative, while trying to show that the justification is indeed adequate and non-arbitrary, I will focus on problems which occur for this kind of attempted justifications.

Albert characterizes the third alternative as making appeal to the "revelatory" nature of specific insights. If I understand him correctly, Albert’s objection against this move is that, as long as ultimate justifications are sought, the demand for further justification can legitimately be raised also concerning any claim to have grasped something self-evident that does not require further justification. He writes:

> It does not help to refer to the revelatory character of certain insights; for the investigator must in the end simply decide whether he is prepared to acknowledge certain presumed insights as revelations. (Albert 1985, 39)

As I interpret Albert, this means that the decision to exempt these items of knowledge from the requirement of further justification, and to accept a particular mode of insight as an infallible criterion of truth or adequacy, can itself be questioned. This decision must itself be such that it is not possible to be mistaken about its correctness. That is, the decision to treat something as not needing further justification, must itself be justified, or shown to be beyond the need for justification. Albert says that this cannot be done without at some point making a non-justified decision, thus vitiating the project of ultimate justification. Otherwise, the regress of justifications would be reopened.

I shall try to make this clearer by presenting another treatment of this issue, by Harold I. Brown. Brown discusses the problems that arise for someone who tries to find infallible foundations of knowledge. In his

\(^3\) See Descartes, *Meditations on First Philosophy*, 158-71 (in the third meditation). The example is controversial, and it has been claimed that there is no circularity in Descartes’ justification. If that is so, then my interpretation of Descartes is mistaken. However, my reinterpretation of his argument can nevertheless serve as an example of a circular justification.
book *Rationality* (1988), Brown argues that the attempt to achieve ultimate justifications, and thus secure foundations, has led philosophers to make appeal to some form of intuition, which is appealed to in order to put an end to regresses of justification. In Albert’s terms, this means that they opt for the third horn of the trilemma by appealing to revelatory insights. Brown discusses the views held by philosophers like Plato, Aristotle, Descartes and the empiricists, and claims that they all fail in a similar way when they appeal to intuition as a way of grasping foundations. I will present his discussion of Aristotle, in order to illustrate the point that Brown makes.

Aristotle realized that not everything can be proven by means of logical demonstration, and held that ultimately knowledge must be founded on first principles which cannot be thus proven, but which are grasped instead by a faculty of rational intuition, *nous*. Brown argues that this is unsatisfactory, because even if we were to accept intuition as a source of infallible knowledge, the problem of how to distinguish genuine intuitions from purported but mistaken "intuitions" remains:

> If intuition is to provide a foundation for rational justification, it must yield self-evident propositions. But if the propositions in question cannot be justified on the basis of independent evidence, it is difficult to know how we can mediate competing claims to have achieved self-evidence, and how we distinguish cases in which we mistakenly believe that we have achieved this grasp. Again we are in a bind: if we allow reasons, we reopen the regress our foundational propositions were supposed to block; and if we do not allow reasons, it is difficult to see how claims to self-evidence can be rational. (Brown 1988, 44)

The problem Brown points to, is that if intuition is to yield adequate foundations, one must be able to be certain that a particular intuition really is a genuine rational intuition or revelatory insight, and to distinguish genuine intuitions from subjective experiences of certainty.

---

4 In *Posterior Analytics*, book I, ch. 3, Aristotle argues that not all knowledge can be demonstrative, and the role of intuition is discussed in book II, ch. 19 of the same work.
which are not genuine intuitions.\textsuperscript{5} In regard to perception, a distinction is often made between veridical and illusory perceptions. One can make a corresponding distinction among intuitions. Then the problem is to distinguish veridical intuitions from illusory ones.

If different persons, for example, have incompatible views which they both claim to have an intuitive grasp of, then how does one decide which of the intuitions is the veridical one? If one were to invoke reasons that are not internal to the purported intuitions themselves, the regress of justification would be reopened.

In other words, if intuition is to provide the foundation we require, it must not only certify truths, it must also certify itself. It is not enough that self-evident truths be true, it must not be possible for us to be mistaken about whether we have achieved self-evidence. If we admit second-order doubts, we face again the very threat of an infinite regress that intuition was supposed to end. (Brown 1988, 48)

Since this is so, the claim that a certain standard of rationality, for example, can receive ultimate justification by its being self-evidently adequate, fails. Appeal to intuitive experiences of certitude cannot remove the possibility of legitimate doubt about its adequacy. Since ultimate justification requires that this possibility of doubt is ruled out, the claim to self-evidence and correctness of the intuition — the claim that it is veridical — must also be justified. Therefore, attempts to achieve such justifications by appealing to insights that reveal particular standards as self-justifying or beyond doubt, are bound to fail.

The argument in this section has been directed against attempts to give ultimate justifications of standards. An attempted ultimate justification forces a proponent to choose between the three alternatives of the Münchhausen trilemma, and none of the alternatives leads to the desired result — a justification which establishes with certainty that a standard is adequate (or inescapable).

\textsuperscript{5} Also see Haack 1978, 235-36, where Haack makes parallel remarks about the purported self-evidence of logical propositions.
2.3 Fallibility: the case of logic

I have argued for a fallibilistic view of standards of rational inquiry. Standards cannot be ultimately justified in a way which removes any possibility of doubt about its adequacy, or its status as an inescapable presupposition.

Now, in the case of many examples of standards that have been proposed, to claim that with regard to them we are fallible, and that critical discussion may lead to their being revised or even rejected, appears uncontroversial. The principle of sufficient justification might be an example of such a standard. Another example might be a Popperian standard of theory choice, stating that "of two theories, it is rational to prefer the one with the highest degree of falsifiability (everything else being equal)". I hope that a claim to the effect that such standards are fallible — in that we may be mistaken about whether or not they are adequate — will provoke no disagreement.

Perhaps it is a more controversial claim to state that the same thing is the case with all standards of rational inquiry, and that this includes the rules of deductive logic. If there is anything that is certain, and from which there is no escape, logical principles and inference rules appear to be the most promising candidates.

Nevertheless, there are philosophers who have claimed that fallibilism applies even to logic, and that logical inference rules are revisable.

In his 1951 paper "Two Dogmas of Empiricism", Quine famously wrote that logic is revisable. Logical laws are seen as parts of a system, where the system as a whole is to be adjusted to experience. If the system conflicts with experience, many different revisions are possible, including revisions of logic:

Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws. Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between
such a shift and the shift whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle? (Quine 1980, 43)

Thus according to Quine considerations of simplicity and empirical accuracy can, in principle, come to motivate revisions of logic.\textsuperscript{6}

Harold I. Brown, who holds a fallibilist view of logic, gives some examples of how the principle of excluded middle (which says that every statement is either true or false (A or not-A)) has been criticized.

This principle is rejected in intuitionist logic. The reason is that intuitionists understand truth in terms of provability, and hold that there are statements which cannot be either proven or disproven, and thus are neither true nor false.\textsuperscript{7}

Another line of criticism against this principle, which Brown mentions, comes from Frederic Fitch. Fitch, according to Brown, has tried to develop a logic which does not include the principle of excluded middle.\textsuperscript{8}

His reason for rejecting it is that he hopes that by doing so he can avoid the paradoxes which Russell’s theory of types was designed to solve. Fitch wants to avoid both the paradoxes and the theory of types. The problem about the latter theory, as Fitch sees it, is that it rules out "self-referential claims, while Fitch maintains that the generality sought by philosophers requires self-reference." (Brown 1988, 75) Brown concludes his discussion of Fitch by noting:

What is particularly intriguing here is that there are trade-offs between accepting certain rules of inference and achieving other cognitive goals, and once this is recognized, we can no longer accept the claim that familiar inferences require no justification. (Brown 1988, 75)

If one does not wish to speak in terms of justification, one could put Brown’s point in other terms, and say that the claim that logical rules are beyond criticism and revision or rejection, should itself be rejected.

According to Brown, yet another line of criticism against the principle of excluded middle comes from empirical studies that psychologists have

\textsuperscript{6} For a discussion of changes in Quine’s view of logic, see Andersson 1998.
\textsuperscript{8} Brown refers to Fitch (1952) \textit{Symbolic Logic}. New York: The Ronald Press.
made of concepts, and how people use concepts to classify things. If the principle in question is maintained, it has the following result:

For a given concept, $C$, excluded middle requires that every sentence of the form $'x \text{ is } C'$ must be either true or false. This, in turn, requires that every concept be fully determinate, i.e., that every concept have a complete set of necessary and sufficient conditions. If this were not the case, then there might be an item, $a$, which could not be classified as either $C$ or not-$C$, and the sentence $'a \text{ is } C'$, would be neither true nor false. (Brown 1988, 75)

He continues:

But psychologists studying the ways in which people categorize items, and the ways in which they relate concepts to one another, have generated a substantial body of evidence that is extremely difficult to explain on the hypothesis that all concepts have necessary and sufficient conditions, but which can be explained much more readily on alternative hypotheses. (Brown 1988, 75)

Brown writes that it is reasonable to think that at least some of our concepts are such that they do not have any necessary and sufficient conditions of application, and suggests instead a view of concepts according to which "we will build into our concepts what we need to deal with the situations we have encountered, and that we may encounter items that we have not envisaged, and whose status will be unclear." (Brown 1988, 75) As an example of such a concept, Brown mentions "furniture", and asks whether telephones fall under it. His answer is that the concept of furniture may be such that there is no clear answer to the question: "The concept of furniture is considerably older than the concept of a telephone; there is no reason to assume that an answer to this question must have been built into the earlier concept" (Brown 1988, 76). The moral of this story, according to Brown, is that "our assessment of the acceptability of a principle of logic can be affected by results in a variety of fields" (Brown 1988, 76).
Now, I do not want to defend this particular view of the nature of concepts, or argue that it shows that the principle of excluded middle is invalid. The point I wish to make, with Brown, is just that the results of inquiries *may* provide reasons for criticizing, and possibly rejecting, logical rules — and that we cannot determine in advance, prior to the results of actual enquiries, that no such reasons can be forthcoming.

Finally, there is also the method of counterexamples for criticizing logical inference rules. To this I will return in a later section.

In spite of considerations such as those presented above, some philosophers have held that some of our logical inference rules must be seen as inescapable and hence unrevisable. The idea is that some of those rules would have to be presupposed in any attempt to rationally criticize and revise anything. One philosopher who holds this view is Karl-Otto Apel. His work is of interest at this point, because he also claims that rules of logic (among other things) *can* be ultimately justified — in spite of the Münchhausen trilemma. Apel’s attempt at ultimate justification will be discussed and criticized in the following sections.

2.4 Does Apel avoid the Münchhausen trilemma?
To begin with, Apel claims that the Münchhausen trilemma presented by Albert is based on a misunderstanding of the nature of justification, and that it is the result of what he calls an "abstractive fallacy" (Apel 1987, 276). Apel holds that the trilemma arises because Albert views logical deduction as the only possible mode of justification, and that Albert has been led to this deductive view of justification because he abstracts from the pragmatic dimension of argumentation and justification. The result of this abstraction, Apel argues, is that Albert understands justification only as a matter of deduction of statements or beliefs from other statements or beliefs. On Apel’s interpretation, this is what leads Albert to claim that any appeal to evidence to end chains of justification constitutes an arbitrary, nonjustified decision.9

9 For Apel's interpretation of Albert's trilemma see Apel 1987, 255-57, 259-62, 276-77; and Apel 1980, 263, 268. This interpretation of the trilemma is accepted by Jürgen Habermas, who writes as follows about it: "It arises only if one presupposes a *semantic concept of justification* that is oriented to a deductive relationship between statements and based solely on the concept of logical inference. This deductive concept of justification is obviously too narrow for the exposition of the pragmatic relations between speech acts." (Habermas 1990, 76)
Then Apel proposes a non-deductive mode of justification, which includes the pragmatic dimension of argumentation, and which he claims can provide ultimate justification of certain standards of rationality and moral norms. Apel’s basic idea is that these standards are necessary presuppositions of argumentation, that they are "conditions of the possibility of intersubjectively valid argumentation" (Apel 1987, 276). They are to receive ultimate justification by means of a transcendental-pragmatic reflection, which is to uncover these necessary presuppositions, and establish that it is indeed necessary to accept them if argumentation is to be possible. They are inescapable. The standards that Apel tries to justify in this way include rules of logic (although he does not specify any particular ones), a moral norm which prohibits lying, and a moral norm which demands that those who participate in any form of argumentation must "mutually recognize each other as participants with equal rights in the discussion." (Apel 1980, 259)

Apel’s interpretation of the Münchhausen trilemma is based on a misunderstanding of Albert’s argument. First of all, Albert does not hold that justification can only be deductive. He does, however, express himself in a way which invites such a misinterpretation. Before presenting the trilemma, he discusses the peculiar merit of valid deductive inferences — that they guarantee the transference of truth from premises to conclusion — and then he actually formulates the trilemma in terms of deductive justification (Albert 1985, 16-18). But later on he adds that

the situation is not essentially altered if inferential processes other than those of deductive logic are introduced in order to bring about the foundation regress. Neither the use of inductive procedures of any kind nor recourse to some transcendental deduction can help remedy the situation [.] (Albert 1985, 20)

I shall try to go beyond what Albert says and explain why this is so, and why the trilemma is not restricted to attempts at deductive justifications.

If one tries to give an ultimate justification of, for example, a standard of evaluation, one alternative is to appeal to something besides that standard itself, and try to show by some mode of argument or inference that the standard is justified on the basis of that something. But if the standard is to be ultimately justified, so that the impossibility of legitimate
doubt concerning its adequacy is established, then that something appealed to in the justification, as well as the adequacy of the way it is related to the standard being justified, must themselves be justified. Since this is so, a potential regress of justifications opens up for any attempt to provide ultimate justifications, whatever the source of justification or mode of inference is held to be. One alternative to such a regress, is to resort to a circular justification. Another alternative — the third horn of the trilemma — consists in trying to show, instead, that the standard can be justified without appealing to anything beyond the standard itself, for example that it can be grasped as self-justifying. It is in this way I have formulated my own version of the trilemma. It generalizes Albert’s formulation.

So Apel is mistaken in thinking that the trilemma is connected to a deductive view of justification. He is also mistaken in his claim that Albert’s reason for rejecting as arbitrary any attempt to provide ultimate justification by appealing to epistemic evidence, is based on the fact that it is a nondeductive process. In his discussion of attempts to achieve ultimate justifications by choosing the third alternative in the trilemma — by appealing to revelatory insights — Albert makes it clear that the reason such attempts are inadequate are not that they are not deductive. The reason is instead simply that such attempts fail to meet the requirements that adequate ultimate justifications would have to fulfil. Albert’s argument is that if something is to be shown to be ultimately justified by being grasped as self-evident or self-justifying, and if this justification is to establish with certainty that the truth or adequacy of the object of justification is beyond doubt, then on every occasion the claim to have grasped something in that way can itself be questioned. And therefore, that claim itself requires further justification. (It should be added, that on my interpretation of the Münchhausen trilemma, it arises for anyone who attempts to find ultimate justifications, but that it need not arise for attempts to weaker forms of justification that do not claim to provide certainty.)

Therefore, I conclude that Apel’s — and Habermas’ — claim that the trilemma can be escaped if one uses a nondeductive mode of justification is mistaken. However, that it is called a ”trilemma”, is based on the fallible claim that none of the three alternatives can result in an adequate ultimate justification. That is, whether the situation Albert describes
constitutes a trilemma or not depends on whether or not it is possible for someone facing the three alternatives to achieve an adequate, non-arbitrary ultimate justification. But that is what Apel claims to achieve by means of his transcendental-pragmatic mode of justification, and I shall therefore turn to an examination of this attempt, and see whether Apel succeeds or fails, on his own terms.

2.5 Apel’s attempt at ultimate justification
Apel’s project of ultimate justification is divided into two stages. First he attempts to justify standards of rationality and morality relatively to participation in critical argumentation. This he tries to do by showing that they are necessary and noncriticizable presuppositions of such argumentation, and that anyone who argues must rely on them. But for Apel that is not enough. Such a justification would not be truly ultimate. The standards would be hypothetical — dependent upon a contingent will to participate in argumentation — rather than categorical. So Apel’s second stage consists in arguing that participation in argumentation — and thus accepting the standards — is not optional, but that it is a presupposition for any meaningful thinking and acting. If that could be justified, the standards would have been ultimately justified as categorical, as inescapable for all human beings.

Since my concern is only with rationality in inquiry, and since some form of argumentation can plausibly be seen as part of any inquiry, I shall deal only with the first stage of Apel’s justification, in which he attempts to justify standards in relation to argumentation. Even if Apel’s second step were to fail, he might be able to show that, at least relative to argumentation and inquiry, there are standards which can be ultimately justified as inescapable. I will argue, however, that he fails to do even this. Consequently, his stronger claim to have justified these standards as categorical will be seen to fail, too.

Apel’s claim is that there are certain standards or rules which function as necessary presuppositions for argumentation. Among these are included standards of rational inquiry like rules of logic, but also what we would normally refer to as moral norms. In his overall project, it is actually these moral norms that Apel is most interested in justifying, since they form the basis of the discourse ethics he develops. It should be
remarked, however, that if Apel is right, these moral norms would be presuppositions of any form of rationality, and would therefore really have to be counted as standards of rationality. But in his discussion with the critical rationalists, Apel has primarily focussed on questions concerning the justification of rules of logic, and the tenability of standards proposed by the critical rationalists. I shall therefore limit myself to discussing the matter in terms of standards of rational inquiry.

Apel refers to his mode of justification as "transcendental-pragmatic reflection". The term "transcendental" signals that this reflection is concerned with establishing conditions of possibility of something. In Apel’s case that something is meaningful argumentation. That the reflection is "pragmatic", means that argumentation is considered not only at the level of statements made, but also at the level of performative speech acts in which arguers make such statements. An important part of Apel’s project of justification consists in identifying contradictions between statements, and presuppositions that are connected to the speech acts (on the pragmatic level) in which such statements are made.

The standards that Apel holds to be susceptible of ultimate justification are supposed to have two characteristics: on the one hand, it should not be possible to justify them deductively. On the other hand, they should be such that it is not possible to challenge them (in argumentation) without being trapped in a performative contradiction. The latter concept, which Apel uses, Habermas explains as follows: "A performative contradiction occurs when a constative speech act 'kp' rests on noncontingent presuppositions whose propositional content contradicts the asserted proposition 'p'." (Habermas 1990, 77) An example of such a contradiction would be if I asserted the statement ("p") "the logical rule 'R' is not valid", and if the speech act ("kp") I performed in making the assertion contained a claim to be able to defend the statement by arguing further for it, and if this further arguing would in turn depend upon my using the very rule 'R' that I purported to reject. Other examples might be the statements "I do not exist", and "You do not exist". Apel’s arguments for ultimate justifications of certain standards depend on identifying such performative contradictions.

The standards that are to be justified have a peculiar status according to Apel. They are "presuppositions of argumentation that one must always (already) have accepted, if the language game of argumentation is to be
meaningful.” (Apel 1987, 277) Their ultimate justification is to be provided by showing that without accepting these standards, meaningful argumentation is not possible. Since that is so, Apel claims, they cannot be criticized or doubted in a meaningful way. If one tries to do that, one will end up in a performative contradiction. That this is really the case, is what Apel intends to establish by means of his transcendental-pragmatic reflection. This reflection, he claims, gives us the insight that certain standards are inescapable, necessary presuppositions which we cannot reject or criticize. This reflective insight, Apel claims, is a sufficient ultimate justification of the standards concerned.

Briefly put, his argument for the ultimate justification of, for example, rules of logic, runs as follows. If someone tries to argue against one of these rules, then transcendental-pragmatic reflection reveals that that person nevertheless presupposes and makes use of that very rule in his argument. Thus, he is guilty of a performative contradiction.

I think the following would be an example of what Apel has in mind: Imagine that I deny that modus ponens is valid, and argue thus: "If a logical rule has the defect $D$ (where $D$ might be that there are cases in which it permits one to derive a false conclusion from true premises), then the rule is invalid. Modus ponens has the defect $D$. Therefore modus ponens is invalid." In delivering this argument, I would be presupposing modus ponens (from "A" and "if $A$, then $B$", derive "$B$"). Apel’s point would be that my statement about the invalidity of the rule is contradicted by what I presuppose on the pragmatic level when I argue for the statement — viz. that the rule is valid.

It should be noted that in this example I argued that the inference rule was invalid because I had, indeed, found a counterexample to it. Even if Apel were to be able to prove that the rule is inescapable, the counterexample would still show that it is not valid. I will discuss this further in section 2.7. But first I will present a general epistemological objection against Apel’s justification procedure.

2.6 A general epistemological objection against Apel’s justification procedure
In order to achieve an ultimate justification of, for example, a rule of logic, it is not enough to show that someone who criticizes the rule has, as
a matter of fact, implicitly been presupposing its validity. One must also show that this is not contingently so, but that it is necessary to presuppose this rule if any argument is to be meaningful and performatively consistent. That is a very strong claim, and I will argue that Apel cannot justify it and that his attempted ultimate justification therefore fails.

To see why Apel fails, I will spell out the conditions that an ultimate justification in Apel’s sense would have to fulfil. Remember, that the justification of a standard consists in showing that meaningful argumentation is not possible without presupposing that standard. In order to establish that this is so, it is not sufficient to show that as a matter of fact someone opposing the standard has presupposed it. What is also required, is that one show that there are no alternatives to that particular standard, which can be used in its place. If, for example, Apel wants to justify some logical rule, it is not enough to show that if we so to speak “remove” that rule from our present set of rules, we are unable to argue. He must also show that there is no other rule with which we could replace the rejected rule so as to restore and maintain the practice of argumentation. That is, if Apel’s attempted ultimate justification of any particular standard is to succeed, he is forced to justify what one might call a “negative existential claim” to the effect that there can be no alternative to the standard in question. How can such a claim be justified, or, how can such a possibility be ruled out? To establish it, Apel would have to show that every possible alternative had been considered, and that none of these alternatives would work. He would have to show that there is a finite set of alternatives and that all of these have been considered. But then the claim that the set of alternatives has a certain limitation, must itself be justified. Either way, he would be unable to justify in an ultimate way a claim to have considered all possible alternatives. Therefore, the claim that there are no alternatives to a particular standard remains exactly — a claim. Therefore, Apel fails to achieve the ultimate justification he sought.

Jürgen Habermas, whose philosophical project is similar to that of Apel’s in that he also tries to justify a discourse ethics on the basis of necessary presuppositions of argumentation, has criticized Apel in a similar way. Habermas rejects Apel’s claim to have achieved ultimate justifications of such presuppositions and of the rules of discourse ethics. His reason for this rejection is his opinion that this justification by means
of transcendental-pragmatic argument must at two crucial points be seen as hypothetical and thus fallible.

According to Habermas, such a justification, which proceeds by showing that to negate certain standards (or discourse rules) results in performative contradictions, "must appeal to the intuitive preunderstanding which every subject competent in speech and action brings to a process of argumentation." (Habermas 1990, 87) Now, Habermas regards this our intuitive knowledge about the presuppositions or rules of argumentation as "in a certain sense not fallible" (Habermas 1990, 95), something he does not argue for. But in order to justify a certain presupposition, which requires showing that denying it leads to a performative contradiction, this implicit knowledge must be explicitly formulated, or reconstructed, in language. And with this linguistic reconstruction, an element of fallibility enters.

But the crucial difficulty for Apel appears at the next stage of the justification, where the proponent who defends a particular rule has to "corroborate, through counterexamples, the proponent's assertion that there are no alternatives to the presuppositions he has made explicit." (Habermas 1990, 94) This is the same kind of objection that I directed against Apel above. Habermas concludes that

the assertion that there is no alternative to a given presupposition, that it is one of the inescapable, i.e., necessary and general presuppositions, has the status of an assumption. Like a lawlike hypothesis (Gesetzeshypothese), it must be checked against individual cases. (Habermas 1990, 94-95)

Therefore, ultimate justifications are not possible according to Habermas, and he therefore regards his own justification of discourse ethics and of standards of rationality as fallible.

Now, if there is anything in logic which is certain or inescapable, then the law of noncontradiction is perhaps the most plausible candidate. I cannot imagine what rational argumentation and inquiry would look like if this principle were to be rejected.

---

10 Habermas probably means that this "intuitive knowledge" is constitutive of the practice of argumentation. But surely, both this practice and the intuitions constituting it may be challenged and revised.
Albert nevertheless argues against the claim — by Apel and Hans Lenk — that the law of noncontradiction must be treated as an unrevisable presupposition of argumentation. Lenk’s argument for this principle, as reported by Albert, is that if it were given up, then any statement whatsoever could be derived from any set of statements. I think, however, that Albert’s reply to the argument sketched has a more general significance, and I shall quote it in full:

This argument presupposes, of course, that the principle is rejected, without any corresponding changes being made of other rules. One would probably have to take quite farreaching changes into consideration, in order to compensate for the rejection of this principle. Someone who argues for it, would be faced by a task for which one would not envy him. But that does not in any way mean that the nonrejectability of the principle could be established in an absolute way, as would be required in order to suspend fallibilism at this point. Neither does it mean that a revision of the idea of criticism based on a transformed logic, could with certainty be excluded. That at present we can hardly imagine what that revision would look like, is certainly true. But to me it seems that our restricted imagination can make no claim to logical dignity. (Albert 1975, 85-86)\(^{11}\)

My conclusion of the arguments in this section is that Apel fails to achieve ultimate justifications. It has actually turned out that he remains trapped in the Münchhausen trilemma drawn up by Albert. His claim to provide ultimate justifications by appealing to an ”insight” obtained through transcendental-pragmatic reflection, has turned out to be yet another inadequate attempt to end a potential regress of justifications by choosing the third horn of the trilemma.

2.7 A counterexample to modus ponens
The idea that some logical principles and inference rules must be viewed as impossible to criticize and revise because they would be presupposed

\(^{11}\) My translation from the German original.
by any rational argument, seems to have an intuitive appeal. It is the basic idea behind Apel's and Habermas' justification procedures. Even philosophers who are generally fallibilists have been attracted by it.

Indeed, even a critical rationalist like Bartley, who holds that rationality is a matter of criticism and that everything can be held open to criticism, holds such a view of logic.\footnote{For a discussion and criticism of Bartley's view of logic, see Andersson 1998.}

He qualifies his statement that everything can be held open to criticism by claiming that there is what he calls a "minimal logic", which is "an absolute presupposition of argument" (Bartley 1984, 134). He adds that this is not incompatible with his critical rationalism, according to which everything can be criticized and should be kept open to such criticism. His claim is instead that "the practice of critical argument and logic are bound together. We can reject logic, but to do so is to reject the practice of argument." (Bartley 1984, 134) So this is not presented as an ultimate justification of minimal logic, since he claims that it and the practice of arguing can as a whole be criticized and possibly rejected. But he does claim that relatively to engagement in argumentation (and therefore, I presume, in inquiry), at least some principles or rules of deductive logic must be seen as unrevisable presuppositions. They are inescapable if we are to be able to criticize anything.

Bartley does not say precisely which principles or rules are included in this minimal logic. The only principle he explicitly mentions is the law of noncontradiction (Bartley 1984, 251).

In *The Last Word* (1997) Thomas Nagel expresses a similar view of logic. Nagel is concerned with arguing against subjectivist interpretations of different forms of thought, including logical thoughts. He holds that the validity of logical thoughts must be seen as independent of our psychological makeup, and that psychological or sociological observations can be used neither to challenge nor to support the validity of logical thoughts. A challenge against logic must according to Nagel take place at the level of logical thought itself.

This refusal, on Nagel's part, to reduce logic and rationality to something merely psychological or sociological, is quite in keeping with how rationality is viewed in the present work: whether a standard is adequate or not is not determined by what we think about it or by psychological facts about us (see especially chapters 1 and 6). Nagel also
states several times that his position does not entail that we are infallible with regard to logic and rationality, or that standards must be seen as unrevisable.\textsuperscript{13}

Nevertheless, Nagel talks about "a framework of methods and forms of thought that reappear whenever we call any specific propositions into question." (Nagel 1997, 69) "Simple logical thoughts", he writes, are such that

there is no intellectual position we can occupy from which it is possible to scrutinize those thoughts without presupposing them. That is why they are exempt from skepticism: They cannot be put into question by an imaginative process that essentially relies on them. (Nagel 1997, 64)

The logical examples Nagel uses, are thoughts about the validity of such general inference rules as modus tollens and modus ponens. About the latter, he writes: "Nothing would permit us to attribute to anyone a disbelief in modus ponens" (Nagel 1997, 77).

I would like to complicate this idea of presupposed logical inference rules, by discussing a counterexample to the rule modus ponens. The counterexample has been presented by Vann McGee.

Modus ponens says that from an indicative conditional like "if A, then B" together with "A", "B" can be derived. If an inference rule is valid, one cannot derive a false conclusion from true premises with its help. If one finds a case where a rule like modus ponens permits the derivation of something false from true premises, one has found a counterexample which shows that the rule is invalid.

Here is McGee's first counterexample:\textsuperscript{14}

Opinion polls taken just before the 1980 election showed the Republican Ronald Reagan decisively ahead of the Democrat Jimmy Carter, with the other Republican in the race, John Anderson, a distant third. Those apprised of the poll results believed, with good reason:

\textsuperscript{13} See, for example, Nagel 1997, 22 and 69.

\textsuperscript{14} McGee presents two more counterexamples in the same paper.
If a Republican wins the election, then if it’s not Reagan who wins it will be Anderson.
A Republican will win the election.
Yet they did not have reason to believe
If it’s not Reagan who wins, it will be Anderson. (McGee 1985, 462)

In this case, McGee comments, "the conclusion of an application of modus ponens is something we do not believe and should not believe, even though the premises are propositions we believe very properly." (McGee 1985, 463)

Perhaps this counterexample, and the others that McGee presents, can be challenged. It seems, however, that it is possible that there could be such counterexamples to rules like modus ponens. I want to discuss what this means for the idea that some logical inference rules must be regarded as inescapable and unrevisable because they would be presupposed in any potential criticism against themselves.

I used the example of modus ponens in section 2.5, to illustrate Apel’s transcendental-pragmatic justification procedure. Apel claims that there are rules which cannot be thus criticized because they would presupposed by the criticism. Note that according to the example, I was making use of modus ponens in my criticism, but that at the same time I had found a counterexample to the rule. For the sake of argument, let us assume that McGee also uses modus ponens when he argues that it is not valid. Where does this leave us?

It shows something interesting, I think. It shows that even if a logical inference rule is inescapable and unrevisable, it can still be invalid and one could find counterexamples to it. A justification which shows that a standard is inescapable, does not show that it is also adequate.

Assuming that one has used, for example, modus ponens in arguing that there is a counterexample that shows modus ponens to be not valid, one would have what might be called an "internal refutation" of the rule.

But if one has found a counterexample to a rule like modus ponens, is one really forced to use the very same rule in one’s criticism? I think not.

For one thing, one can try to construct the argument in a different way, without using the inference rule which is being criticized.
Also, note that modus ponens is a completely general rule. It says that from "A" and "if A, then B" one may derive "B" — whatever "A" and "B" are. If one has found a counterexample to a rule, then one may be able to identify some particular feature of the example which makes it differ from instances of the rule which seem unproblematic. Concerning his own counterexamples, McGee remarks: "Each has as a premise a conditional whose consequent is itself a conditional." (McGee 1985, 464)

Then, one can propose a restricted version of modus ponens which rules out all cases in which the consequent ("B") contains a conditional. The counterexamples do not affect this rule, and one can then try to use this restricted rule in one's criticism of the general rule.

So far I have assumed that in criticizing a logical inference rule, one must appeal to some general inference rules which the argument instantiates. But is this really necessary? Is it not sufficient if one can recognize each particular step in the argument as valid — without knowing which general rule such a step is an instance of? It seems reasonable to accept that this could be rational, since in general people can recognize the validity or invalidity of particular arguments without knowing anything about the general inference rules the arguments instantiate.

It may be the case that a particular inference is valid only if it falls under a valid general inference rule. But this is, one could say, a question about the metaphysics of logic. Epistemologically, on the other hand, it seems to be the case that it is actually more difficult to tell if a general inference rule is valid, than if a particular inference is. To require that in order for a particular inference to be rational one must be able to justify it by showing that it is an instance of a valid general rule, seems misguided from an epistemological point of view.

It might be objected, perhaps from a critical rationalist, that in a rational argument one must make use of general inference rules. The reason would be that otherwise one would not be able to criticize arguments by showing that they are of an invalid form — that they are instances of an invalid general rule. But this seems equally misguided, because every argument, whether valid or invalid, is an instance of some

15 Cp. Russell, who writes that "in all cases of general principles, particular instances, dealing with familiar things, are more evident than the general principle." (Russell 1959, 112)
invalid inference rule. Indeed, every valid inference rule falls under some more general, invalid rule. For example, every instance of the modus ponens schema "From 'A' and 'if A, then B', derive 'B'" is an instance of the following schema which is obviously invalid: "From 'A' and 'B', derive 'C'".

Thus, it appears that a particular argument may be rational without making use of general inference rules.

What the considerations in this section show, is that it is not so easy to determine precisely what one must presuppose about logic as one criticizes logic. If one has discovered a counterexample to a logical inference rule but recognizes that one has made use of the same rule in the critical argument, several options are open: one can try to use some other rules, or propose a restricted version of the rule which avoids the counterexample, or judge that every particular inference in the argument is valid.

Thus, it appears that we can avoid treating logical inference rules as inescapable and unrevisable presuppositions of rationality. This further undermines Apel’s attempt to give ultimate justifications of standards.

2.8 Apel’s argument against fallibilism
In concluding this chapter, I shall examine Apel’s claim to have shown that any position which (like critical rationalism) holds that we are fallible with regard to all standards, is untenable. Of course, if he had succeeded in achieving an adequate ultimate justification, then that would by itself have been a sufficient refutation of fallibilism. But I shall now consider an argument that is directed against the claim that we are fallible with regard to all standards.

Apel’s objection is that fallibilism is pragmatically inconsistent, because any argument for it leads to a performative contradiction. Curiously enough, Habermas seems to endorse this objection, in spite of his rejection of ultimate justifications. He reconstructs Apel’s argument, and I will consider Habermas’ version of it. I believe the argument fails, but it provides an opportunity to consider some possible misunderstandings of fallibilism (and critical rationalism).

Habermas asks us to consider an opponent of Apel’s position, who has introduced the Münchhausen trilemma, and then draws the conclusion
"that attempts to ground the universal validity of principles are meaningless. This the opponent calls the principle of fallibilism (f).’’ (Habermas 1990, 78) Now, I do not understand why Habermas uses the term "meaningless" here. It is not clear, either, what is to be meant by "grounding" principles. For a fallibilist, the relevant sense of grounding would be that of providing an ultimate justification, since that would be inconsistent with fallibilism. I would therefore like to reformulate the principle (f) so that it reads "ultimate justifications of the validity (or adequacy, or inescapability) of principles are impossible". If Habermas and Apel are out to get the rationality fallibilists (here represented by the critical rationalists), I think they ought to agree to this reformulation of (f), since if their attack is to be successful they should be careful not to misrepresent their opponents’ position. Then Habermas argues that in arguing for this, the fallibilist opponent must rely on presuppositions of argumentation, which contradict the content of (f):

This is in fact the case, since the opponent, in putting forward his objection, necessarily assumes the validity of at least those logical rules that are irreplaceable if we are to understand his argument as a refutation. Even the consistent fallibilist has, in taking part in a process of reasoning, already accepted as valid a minimum number of unavoidable rules of criticism. This state of affairs, however, is incompatible with (f). (Habermas 1990, 78)

This argument fails, for several reasons. First of all, the fallibilist does claim that with regard to all standards, including the rules of logic mentioned, we are fallible. But he does not claim that they are not valid. Therefore, to accept standards like the logical rules in question as valid, is not at all incompatible with (f). An incompatibility would result, if the fallibilist held a further standard to the effect that it is impermissible to accept or use any rule that is fallible — which would be rather crippling for a fallibilist. But that is not the case with Albert or Bartley, or any other rationality fallibilist that I know of.

Furthermore, rationality fallibilists like the critical rationalists would grant that in every argumentation situation one has to rely on some
rules,¹⁶ but that does not mean that there is any set of rules which must be relied on in every argumentation situation — which would be one requirement for convicting the fallibilist of performative selfcontradiction. Actually, a fallibilist might well agree that there may actually be such a set of rules. On the other hand, if we could know with absolute certainty that there are such rules, fallibilism would be refuted. But to refute fallibilism on this point, one would have to give an ultimate justification which establishes with certainty that these rules cannot be replaced or revised. As has already been shown, this would require that Apel and Habermas could prove that there are no alternative rules which could be used instead — which they are unable to do.

So I conclude that Apel’s and Habermas’ argument against a fallibilistic view of standards fails, and that such a view can indeed be consistently defended.

However, in order to avoid being misunderstood, I want to emphasize that my arguments in this chapter only concern Apel’s claims about ultimate justifications. Apel’s mode of argumentation, where one brings in the question of pragmatic presuppositions, may be very fruitful. It may be one important way of trying to resolve conflicts about the adequacy of particular standards of rationality. It is just that it does not help us to achieve ultimate justifications.

2.9 Conclusion
We are fallible with regard to all standards of rational inquiry, including logical inference rules. Standards cannot be ultimately justified either as adequate or as inescapable. Every attempt at ultimate justification leads to the Münchhausen trilemma, and none of the alternatives of the trilemma can result in an ultimate justification. This might be described as a fallible impossibility result.

Karl-Otto Apel’s attempt to give ultimate justifications of standards by means of transcendental-pragmatic reflection fails, and contrary to his own assertions he remains trapped in the Münchhausen trilemma.

¹⁶ See Albert 1985, 55-56, and Bartley, who emphasizes that he does not ”suggest that all assumptions be abandoned, only that they be held open to criticism!” (Bartley 1984, 121, n. 12)
The idea behind Apel’s justification procedure, that certain standards can be shown to be inescapable presuppositions of rationality because any argument against them would presuppose them, has been criticized. It is not so easy to determine what is presupposed by an argument against, for example, a logical inference rule. I have argued that such arguments can make do with very weak assumptions about which inference rules are valid. I also argued that an agent who appears to be trapped in what Apel calls a "performative contradiction" when he criticizes a logical rule, can avoid such a contradiction by restricting the logical assumptions he uses in his criticism.

Still, philosophers have argued that although strictly speaking we are fallible with regard to standards of rational inquiry, since there are no ultimate justifications, there must nevertheless be some standards which cannot be rationally evaluated and revised. These arguments appeal to claims about the nature of rational evaluations and rational justification. It is arguments of this kind that will be discussed in the next chapter.
3

A CRITIQUE OF ARGUMENTS FOR UNREVISABILITY

3.1 Introduction
In the previous chapter, I argued that we are fallible with regard to all standards of rational inquiry, including evaluative standards, methodological rules and logical inference rules. There are no ultimate justifications of standards, so we have no way of establishing with absolute certainty that our standards are fully adequate, or that they are inescapable presuppositions of rationality.

In this and the next chapter I will examine arguments that have been put forward in favour of a weaker claim. This is the claim that, although we are strictly speaking fallible with regard to standards, there are nevertheless some standards of rational inquiry that are rationally unrevisable. Unlike attempts at ultimate justifications, these arguments do not aim at establishing a conclusion with absolute certainty. More modestly, they are arguments to the effect that given certain reasonable assumptions about the very nature of rational evaluation (or justification, or change) it follows that there must be some standards that it is impossible to rationally evaluate and revise.

As I have already stated, all parties to the discussion agree that some standards can be rationally evaluated and revised, by drawing on some further standards in terms of which such changes are evaluated. The dispute instead concerns whether or not there is a certain subset of standards that cannot themselves be rationally evaluated and revised. Some philosophers argue that there must indeed be such unrevisable standards.

The arguments I discuss here are to the effect that rationality is in a certain way limited. They are supposed to show that there are some standards that enable us rationally to evaluate and change other things —
including other standards — but which cannot themselves be subjected to rational evaluation and rational change: with these standards we supposedly reach the limits of what can be subjected to rational scrutiny.

The arguments examined in this chapter are supposed to show that for reasons having to do with the nature of rational evaluation, there must be some standards which are such that we need them in order to engage in rational evaluations in the first place, but which cannot themselves be made the subject of such evaluation. Since it is only possible to rationally revise a standard if the standards involved in the revision — the standards before and after the change — can be evaluated and judgments about improvement made, the arguments would if correct also show that the standards in question are rationally unrevisable.\(^1\)

At this point I think it is important to distinguish between truth-claims and normative policies or proposals. It is one thing to claim that there are some standards that cannot be rationally revised. This claim I will refer to as the "unrevisability thesis". The thesis states something which can be true or false. It is another thing to propose or adopt a normative policy which says that some standards ought to be treated as unrevisable. That I will refer to as the "unrevisability policy". The unrevisability policy says that some standards always ought to be treated as unrevisable, come what may.

In contrast, a "revisability thesis" would say that there is no standard which cannot be rationally revised: every standard is revisable. That is either true or false. There is also a related normative policy, according to which all standards ought to be treated as revisable: according to this "revisability policy" it is not rational to exempt any standard from possible criticism and treat it as unrevisable.

One can discuss whether these theses are true or not, and one can also discuss whether the policies ought to be adopted or not. But how are the theses and policies related? If the unrevisability thesis were true and some standards were indeed unrevisable, it does not follow that the corresponding unrevisability policy must be adopted.\(^2\) But if an argument succeeded in

---

1. Revisability presupposes evaluability, but I do not claim that evaluability presupposes revisability.
2. Particularly if the individual standards that are unrevisable cannot be identified and singled out.
showing that the unrevisability thesis is true, then this could plausibly be seen as telling in favour of the unrevisability policy and against the revisability policy.

My aim in this chapter is to criticize arguments that have been used to support the unrevisability thesis. My claim will be that we do not have good reasons to accept the thesis that there must be some standards that are rationally unrevisable. The arguments are, one could say, direct arguments for unrevisability. They are supposed to take us directly from claims about the nature of rational evaluation or rational justification to the conclusion that necessarily there are unrevisable standards. Since these arguments do not work, I will conclude that we do not have to accept the unrevisability thesis.

In the next chapter I will examine an indirect argument for the unrevisability policy and the unrevisability thesis. The indirect argument has it that the revisability policy of treating all standards as open to critical evaluation and rational revision leads to a form of relativism which is unacceptable, and that the only alternative is to adopt the unrevisability policy and treat some standards as unrevisable. This argument is criticized in the next chapter.

I shall not attempt to argue in favour of the revisability thesis. In this chapter I shall instead argue that we do not have good reasons for the unrevisability thesis, and that we are not forced to accept the claim that there are any unrevisable standards. And in the next chapter I first criticize the claim that the revisability policy leads to relativism, and I argue that if relativism is to be avoided it is on the contrary preferable to adopt the revisability policy and refuse to treat any standards as immune from critical evaluation and revision. Thus the outcome of chapters 3 and 4 will be that we are not forced to believe that there are unrevisable standards, and that it is rational to treat all standards as rationally revisable.

But now I turn to the arguments that have been used to support the unrevisability thesis.
3.2 Hauptli on the limits of evaluation

In this section, I shall examine and criticize what I call the "limits of evaluation argument" for the unrevisability thesis. The conclusion of the argument is that some standards of rational evaluation must themselves be beyond the possibility of rational evaluation. Since rational revision requires rational evaluation, it is also an argument for the unrevisability thesis.

Bruce Hauptli presents this argument in his *The Reasonableness of Reason*. He uses it as an argument in favour of a position he calls "naturalism". He opposes a position he calls "justificatory rationalism" that demands that everything we accept, including standards of justification and evaluation, must be rationally justified (the principle of sufficient justification). Hauptli is himself a kind of justificationist, and holds that "we should accept only those beliefs, theories, and commitments which we can rationally justify." (Hauptli 1995, 69)

But the requirement for justification does not apply to some things according to Hauptli’s naturalism. In particular, he claims that it does not apply to the justification requirement itself, nor to other standards of rational evaluation. His argument for this draws on the nature of evaluations:

Naturalists maintain that the justificatory responsibility does not apply in this case because of the *special role* these standards play in our justificatory and evaluative endeavors: they provide the presumed background against which our justificatory and evaluative endeavors take place; and, thus, they cannot be subject to such evaluation or justification. (Hauptli 1995, 69)

Here, Hauptli writes about both justification and evaluation. I believe that this argument is more general than the limits of justification argument (to be examined in the next section), and that it can be spelt out in terms of evaluations alone. I shall also assume that when Hauptli speaks about

---

3 Hauptli distinguishes between two kinds of naturalism: one is of a Wittgensteinian kind and is called "therapeutic naturalism", while the other is called "explanatory naturalism" and is associated with Quine. It appears that the thesis of the limits of evaluation is taken by Hauptli to be part of both these kinds of naturalism, but I shall not discuss whether it would be appropriate to attribute it to Wittgenstein and Quine.
justification and evaluation here, what he means are evaluations in terms of what I call "standards of rational inquiry" (see chapter 1). I take it that he is not including for example ethical or aesthetical standards and evaluations.

I shall also assume that Hauptli does not mean that all standards are such that they cannot be evaluated. It is more likely that he is thinking of some subset of standards which are more basic than others, and that those basic, unevaluable standards can be used to evaluate other standards. How large this set of unevaluable standards is supposed to be is not clear from the text, however.

I believe that the following reformulation, which I will call the "limits of evaluation argument", captures what Haupli wants to say. It goes as follows:

In every rational evaluation, one appeals to and thus presupposes standards of evaluation. Since every rational evaluation presupposes standards of evaluation, these standards cannot themselves be evaluated (they are what makes evaluations possible in the first place).

Corollary: Since rational revision presupposes rational evaluation, these standards are also unrevisable.

It is true that whenever we evaluate something, we at least implicitly appeal to and presuppose some evaluative standard or set of standards (and usually a whole lot more besides evaluative standards). It is not possible to assume an evaluative standpoint without at the same time assuming that some standards are adequate, so we cannot evaluate all our evaluative standards at once. But from this it does not follow that there are standards which are beyond evaluation.

According to the argument, standards are presupposed in every evaluation. This may be interpreted as meaning either that in every case some of our standards (of rational inquiry) are presupposed, or that all of them are.

It is true that at least some of our standards are presupposed in any given evaluation. But if only some of them are appealed to in a given situation, then this leaves open the possibility that they can be used to evaluate some of
the standards which are not appealed to at that point. So if the argument is to be correct, ”standards” must be taken to mean ”all our standards”.

But this is simply false. At every given stage of inquiry, there will always be a whole set of different operative evaluative standards. Such a set is not a monolithic package. Particular standards are used for evaluating something in a particular respect, and not all evaluations are evaluations of something in all respects. We will have, for example, logical inference rules, consistency requirements and other evaluative standards, heuristic methodological rules, general standards such as the principle of sufficient justification or perhaps a criticist alternative to it, meta-standards, and so on. Typically, only a subset of our operative standards will be relevant to (presupposed by) a particular evaluation.

Furthermore, even if it were granted, for the sake of argument, that in every case of evaluation all our operative standards are involved, the sought conclusion would still not follow. At least some standards can be criticized (and thus evaluated) by being applied to themselves (for example, a standard saying that everything one accepts rationally must be held open to criticism may be evaluated by examining whether it can itself be held open to criticism; the standard saying that everything one accepts must be sufficiently justified may be evaluated by seeing whether this standard is itself sufficiently justified). If the outcome of such an evaluation by self-application is negative, it constitutes a criticism of the standard although it is itself appealed to in this evaluation (see the discussion of internal refutations of logical inference rules in section 2.7).

Similarly, some evaluative standard or standards may be used to evaluate the whole set of standards of which it is itself a part (is it consistent? is it coherent? is it suitably related to the goals of the inquiry?). In such cases, although a standard is presupposed in an evaluation, it can be used to evaluate either itself or a set of standards of which it is a member.

Hauptli’s claim that it is impossible to evaluate evaluative standards turns in an important way on the thought that standards cannot be applied to themselves. This he tries to support by a discussion of the standard metre in Paris, and the role it plays (actually, played) in measurements of length. He says that the example of the standard metre illustrates that ”a standard cannot
be applied against itself in the way it can be applied against those things it functions as the standard of since it provides the criterion for such an evaluation.” (Hauptli 1995, 71)

Hauptli’s analogy does not work very well, since there are crucial differences between standards of measurement and standards of rational inquiry. The case of the standard metre is significantly different from the case of standards of rational evaluation. Hauptli says that the point of the standard metre, was that it was used to define a length unit. Therefore, applying it to itself to measure its length is, Hauptli remarks, pointless because “the standard and the object to be ‘evaluated’ are one and the same.” (Hauptli 1995, 73) Since the standard metre was used to define ”one metre”, the question as to how long it is does not arise, and as long as it is used as a standard of measurement, there is no other standard which is relevant to its measurement (its ”evaluation” in terms of length).4

But this is not the case with standards of rational evaluation. There are at least three striking differences: (i) Rationality is a normative concept, and standards of rational evaluation do not define what ”rational evaluation” means or is. For example, a standard saying that sets of statements ought to be consistent to be rationally acceptable does not define ”rationally acceptable”. It is always intelligible to ask, concerning any proposed standard of rational evaluation, whether it is really rational to use it (see section 1.4).

(ii) Whereas the standard metre was, for a time, the unique standard for a particular unit of length, this is not the case with standards of rational evaluation. At any given stage of inquiry, there are always a number of different standards which all fulfil the function of rationally evaluating things, but in different respects — standards which may thus be used to evaluate each other. Even if it were the case that no standard of evaluation could be applied to itself, it would not follow from this that no standard of evaluation can be evaluated — unless one were to assume that there is only one evaluative standard.

(iii) At least some standards can indeed be applied to themselves, in a way the standard metre cannot.

4 For a quite different view of the case of the standard metre, see Kripke 1980, 54-56 and 75-76.
Apart from these disanalogies between the cases, there are also complications with the standard metre example itself. For one thing, even if one were to grant that the standard metre could not be measured as long as it was used to define the length unit "one metre", it was indeed possible to evaluate it in other respects. To begin with, the material object which was to serve as standard metre had to have properties that were suitable for a standard of length. Not any old object would do. For obvious reasons, it had to be made of a material such that the length of the bar would not be significantly affected by changes in temperature or pressure, by corrosion or radioactive decay, and so forth. In order that the bar should be able to function as a standard of length ought to, it should be a very specific type of object, and in selecting a particular material, physical theory and information about the properties of different materials were used.

Both standards of rational inquiry and standards of length are something we use, and how good they are as standards (whether for evaluation or for measurement) may depend on changing characteristics of the situations in which we put them to use.

Interestingly, as Hauptli himself mentions later on, "one metre" eventually came to be defined instead "as 1,650,763.73 vacuum wavelengths of the orange-red radiation of Krypton 86 under certain specified conditions." (Hauptli 1995, 125) He also gives an explanation of why this new standard replaced the older one: developments in microphysics gave rise to a need for exact measurements of very short lengths. These were "most easily measured in terms of wavelengths", and "the accuracy of this method of measurement so surpassed those which applied to the lines on the bar that the new standard came to be adopted." (Hauptli 1995, 125)

What this illustrates is that standards of measurement are supposed to fulfil particular functions, besides that of just defining units of length. We want to be able to use them in certain ways, and they can indeed be evaluated by examining how well they can be so used. It also illustrates that in the course of inquiries new developments may require such standards to fulfil new functions and thus motivate revisions of the standards. In this, they are similar to standards of rational evaluation.
Hauptli’s attempt to argue that some standards must be unevaluable is mistaken. To summarize, I have three central objections against it. (1) Hauptli treats our evaluative standards as though they are one monolithic package. That is actually required if his argument is to work. But instead, for any inquiring agent, there is always a vast number of standards of different kinds. Thus it might be possible to set one standard (or a limited set of standards) aside for a moment, so to speak, and use some other standard to evaluate the first one. (2) While it is true that not all standards can be applied to themselves, some standards are self-applicable. This opens up still further possibilities for evaluating standards. (3) Hauptli’s analogy between evaluative standards and the example of the standard metre, which was supposed to show that standards cannot be applied to themselves, is highly misleading and shows no such thing.

My conclusion is that the limits of evaluation argument fails to show that some standards of rational evaluation are beyond the possibility of rational evaluation and revision. Therefore, neither does it support the normative policy according to which we ought to (or must) treat some evaluative standards as beyond evaluation and revision.

There is however a similar but distinct argument that has been held to show that there are such limits to rational evaluation and revision, an argument which appeals to the nature of rational justification.

3.3 The limits of justification argument
The limits of justification argument for the unrevisability thesis is, one could say, a by-product of the classical problem of justifying rationality. It is a regress argument which is based on certain assumptions about rationality and about the connection between rationality and justification.5 The alleged

---

5 This argument is distinct from, but related to Albert’s argument against ultimate justifications presented in the previous chapter. That argument was to the effect that neither of the alternatives in the Münchhausen trilemma — infinite regress, circularity or appeals to something like self-evidence — can result in ultimate justification. It is also distinct from the standard ”regress argument” which has been used to defend a foundationalist account of justified empirical belief. That argument focuses only on beliefs, and has been used to support the conclusion that if any empirical beliefs are justified, then there must be some
conclusion is that there are a number of standards of rational justification (a kind of evaluative standards) which are not susceptible of evaluation. Therefore, they are also unrevisable.

The problem has to do with how to justify the standards of justification themselves.

The first premiss is the "principle of sufficient justification" (PSJ). It can be formulated as follows: "It is rational to accept something if, and only if, it is sufficiently justified". Another assumption is that in order to justify something, one must appeal to some standard of justification (just as, more generally, every evaluation requires appeal to some evaluative standard). A third and crucial premiss is that it is, indeed, possible to be rational and to justify things.

The PSJ says that in order for us to accept something we must be able to justify it, and that if something is justified, then we ought to accept it. But what about the standards of justification that we appeal to as we justify things, and what about the PSJ itself? According to the PSJ, our acceptance and use of these standards is rational only if we can justify them (including the PSJ).

Here the trouble starts. In order to justify the PSJ or some other standard, we would have to appeal to some further standard. It would have to be a different standard than the one being justified, in order to avoid begging the question. At this point, an infinite regress threatens to open up: if we justify standard $S_1$ by appeal to (among other things) standard $S_2$, the PSJ requires that we justify $S_2$. If we justify $S_2$ by appealing to $S_3$, then $S_3$ in its turn must be justified, and so on.

foundation or basic beliefs that are justified independently of other beliefs (for discussion see, e.g., Haack 1993, 21-25; Everitt and Fisher 1995, 70-72).

It can be compared with the arguments, which do focus on the justification of standards, in Ayer 1956, Wedgwood 1999 and Brown 1988.

The limits of justification argument largely overlaps with Bartley's account of the failure of panrationalism, and the way in which the recognition of this failure is incorporated into the kind of position he refers to as "critical rationalism" (Bartley 1984, chh. 4-5). Bartley, however, goes on to criticize this kind of position for its relativistic consequences, and argues that justificationism in general should be rejected in favour of a criticist theory of rationality. Here, I want to examine only the claim that this argument supports the unrevisability thesis.

It is noteworthy that this argument, as Bartley points out, is used by both skeptics, fideists, and relativists.
If we accept that rationality requires justification, and if we also insist that rationality and justification is possible, then we must put a stop to this regress at some point.

The only way to do this, the argument continues, is to accept that rationality necessarily has its limitations: we have to accept the PSJ and any other standards that we need to be able to justify things — without justification. These standards would be the basic standards, on which the possibility of justifying other things rests. Since the acceptance of these basic standards which are to terminate the regress is unjustified, it is according to the PSJ irrational or, perhaps, nonrational. According to John Worrall, at this point "the only option is [...] the honest admission that ultimately we must stop arguing and 'dogmatically' assert certain basic principles of rationality." (Worrall 1989, 383)

Alternatively, one can follow Alston in distinguishing between on the one hand showing that one is justified in accepting something, and on the other hand merely being justified in accepting something (Alston 1976). Alston proposes that one might be justified in accepting something, without having a justification for it — that is, without being able to appeal to standards and show that it is justified. Then one could perhaps be justified in accepting the basic standards, without being able to show that the acceptance is justified. In that case, one’s accepting the basic standards would not automatically be nonrational or irrational (according to the PSJ). But it would still be the case that the basic standards could not be shown to be justified (or unjustified) by appeal to any standard of justification.

What I want to consider is the next step in the argument, which is supposed to take us from the need for nonrational adoption of standards of

---

6 Accepting only the PSJ itself would get us nowhere; it does not specify what counts as rational justification. Therefore, some further standards of justification are needed.
7 Some justificationist philosophers, who accept the PSJ, maintain that rationality, the PSJ, and other evaluative standards can indeed be rationally justified. See, for example, my discussion of Apel and Habermas in chapter 2. Also, see Siegel 1997, ch. 5. They all argue that rationality and standards of rational justification can be justified in a self-reflexive way without begging the question. But this would mean, then, that these standards are open to evaluation (Siegel explicitly claims that standards are also revisable, and I believe that Habermas would also accept this).
8 For a criticism of Alston’s attempt to use the concept of being justified for answering the skeptic’s regress argument, see Hauptli 1995, ch.2.
justification, to the claim that these standards, once adopted, cannot be rationally evaluated.

The argument would continue: Given these basic standards, we can justify other things, for example other evaluative standards that can then be used in rational evaluations. The "basic" standards are themselves beyond rational evaluation, however. We cannot argue for or against them, since any rational argument presupposes them (or some other evaluative standard which can be justified on the basis of them). And since they cannot be rationally evaluated, they cannot be rationally revised. So the argument goes.

But I think that this is incorrect.

First of all, we must stop at some point by accepting some standard without justification. But it does not say that there is some point at which we must stop. Thus, even if we have decided to terminate the regress at a particular point by adopting a certain standard, we could perhaps change our minds and decide instead to adopt some other standard further down the regressing chain of justifications. If we do, the standards we come to adopt instead could be used for justifying, or criticizing, the standards we had previously stopped at. At least this is not ruled out by the argument.

But even if one were to grant that there is some point at which we must terminate the regress of justification, it would still not follow that the standards we adopt nonrationally cannot be evaluated.

Assume that we have nonrationally adopted a particular set of standards. This set includes the PSJ. The argument above is supposed to show that these standards cannot be evaluated. But I see no reason to think that, given that we have come to accept these standards, we are necessarily unable to evaluate them.

Our set of standards contains the PSJ, for example. When we apply it to itself (or to any of the other unjustified standards we have adopted), it tells us that it is not rational of us to accept it (since it is not justified). That is certainly an evaluation.

Similarly, if the set contains a consistency requirement, we can apply this to the set of standards (or a subset of it) and evaluate it (or a subset of it) in terms of internal consistency. Once again, this is a way of evaluating the evaluative standards. And I do not see any reason to suppose that the set of
standards we have adopted (or their further applications) are guaranteed to be consistent. By hypothesis, we did not perform any check for consistency before we committed ourselves to them.

Furthermore, once we have a first set of standards of justification in place, we may come to make justified additions to that original set of standards. Is it not possible that some new proposed standard will be judged to be justified on the basis of some of the original standards, and added to the set? Then it is possible that new ways of evaluating the standards we have hitherto possessed become available, as we add new standards to those we already accept.

To this one might object and say that standards that have been added at later stages cannot be used to criticize and revise the standards introduced at earlier stages and on the basis of which the new additions are judged to be rational. If, for example, one of my original standards \( S_1 \) allows me later on to add a new standard \( S_{47} \), then \( S_{47} \) cannot be turned against \( S_1 \) and used to criticize and revise it.

The objection is mistaken, because in some cases exactly this can take place. The following example shows this. Assume that one of my original standards is the meta-standard \( M_S \), according to which a standard is rationally acceptable if it is reliable. A reliable standard produces correct results often enough, but it need not be guaranteed to do so every time. Besides \( M_S \), I have other standards, such as logical inference rules and various evaluative standards, all of which are reliable in this sense. Then I go through a sort of Cartesian meditation, and discover a new standard \( D_S \). According to \( D_S \), it is rational to accept only what I perceive clearly and distinctly to be correct. I believe that I can use \( D_S \) to discriminate among beliefs and among standards, and that it is guaranteed to lead to the right results. It is thus a reliable standard, and therefore rationally acceptable according to the original standard \( M_S \). So I adopt \( D_S \). But once I have done so, \( D_S \) can be used to criticize and reject \( M_S \). The \( M_S \) allows me to accept standards that are merely reliable (and to accept beliefs on the basis of reliable standards), but according to \( D_S \) this is too permissive, so the meta-standard \( M_S \) must be rejected as stating a too weak sufficient condition for what is rationally acceptable.
Perhaps it is unrealistic to take a standard like DS as example, but the general structure presented here is possible. A different example might be formulated in terms of a quite general form of modus ponens and a form in which the rule has been restricted in some way so as to exclude counterexamples (as discussed in the previous chapter). So, the possibility that standards which later on come to be added to an original set of standards can be used to evaluate and revise the original members of the set, cannot be ruled out.

If these remarks are correct, then from the claim that standards of evaluation must be nonrationally adopted, it does not follow that they cannot be evaluated. And the thesis that they must be unrevisable, does not follow, either.

3.4 Conclusion
In this chapter I have discussed two direct arguments for the unrevisability thesis. Both arguments were supposed to show that there must be some standards which are such that while they enable us to evaluate other things, they cannot themselves be rationally evaluated (they are unevaluable). Since a rational revision of standards presupposes that the standards can be rationally evaluated, it would also follow that the standards in question are rationally unrevisable.

I have criticized these arguments and argued that they both fail. They do not force us to accept the unrevisability thesis, according to which there are standards which cannot be rationally revised. I do not claim to have shown, or even to be able to show, that the thesis is false. That is, my own arguments do not show that the revisability thesis is true and that all standards can be rationally revised. What I shall later defend is instead the revisability policy — the normative policy according to which no standard ought to be treated as immune from criticism and revision. That policy can be defended without first establishing the related revisability thesis.

But the revisability policy has been challenged and argued to lead to an unacceptable form of rationality relativism. A successful argument to that effect would be an indirect argument for the policy of treating some
standards as rationally un revisable, and I admit that it might be taken to support the un revisability thesis as well. Such indirect arguments for un revisability is the topic of the next chapter.
4

A CRITIQUE OF ARGUMENTS AGAINST REVISABILITY

4.1 Introduction

In the previous chapter I argued that the direct arguments for unrevisability, which purport to take us straight from claims about the nature of evaluation or justification to the unrevisability thesis, do not work. In this chapter I examine indirect arguments for unrevisability. These arguments are to the effect that to treat all standards as revisable has unacceptable relativistic consequences, and that therefore some version of the unrevisability thesis or policy must be accepted.

I argue that these indirect arguments fail as well. They fail to show that the revisability thesis or the revisability policy have any relativistic consequences. On the contrary, I suggest, it is the policy of treating standards as exempted from critical evaluation (and revision) which leads to the relativistic conclusion. The rational option, I will propose, is actually to adopt the revisability policy and refrain from treating any standards as immune to criticism and revision.

The objections against revisability that I discuss in this chapter have been brought forward by John Worrall. His formulation of them is not very detailed, and I shall try to bring out the assumptions that are implicit in his objections against the thesis and policy of revisability.

In Science and Values (1984), Larry Laudan proposes an account of scientific change he calls "the reticulated model". Laudan claims that not only factual beliefs, but also scientific methods (including standards of rational evaluation) and the goals of scientific inquiries are open to change. He also claims that none of these elements need to be treated as fixed, and that through a rational process of gradual, piecemeal change, any belief,
standard or goal may come to be rationally revised. Laudan’s reticulated model will be discussed in chapter 5, in relation to the so called "bootstrap view" of rationality.

In the review article "The Value of a Fixed Methodology" (1988), John Worrall criticizes Laudan’s model and several of the claims he makes about scientific rationality and change. Laudan responded to some of the criticisms in "If it Ain’t Broke, Don’t Fix it" (1989). Worrall repeated and extended his objections in "Fix it and be Damned: A Reply to Laudan" (1989). One of the things Laudan and Worrall disagree about, is to what extent the methods of science have actually changed, and how particular changes in the history of science should be described. I will not comment upon that part of their discussion. Worrall also criticizes Laudan’s interpretation of certain other philosophers of science, and objects to particular details of Laudan’s reticulated model. I shall not go into that either.

In this chapter, I will instead examine two objections that Worrall directs against the revisability policy and against the claim that all standards are revisable. The objections occur in his exchange with Laudan.

However, first I have to say that from what Worrall writes, it is not quite clear if his objections are directed against the revisability thesis or against the revisability policy — or against both of them at once. They might be interpreted as criticisms of the thesis, the claim that all standards are revisable. Thus interpreted, the objections are part of an indirect argument of the following form:

Assume the revisability thesis. If the revisability thesis is true, it follows that rationality relativism is true. But rationality relativism is false. Therefore the revisability thesis is false, and its negation, the un-revisability thesis is true.

I will try to argue that his objections do not show that rationality relativism follows from the revisability thesis, and thus that these indirect arguments for the unrevisability thesis do not succeed.

But Worrall’s objections can also be interpreted as attacks on the normative revisability policy, and thus instead as indirect arguments for the
unrevisability policy according to which we ought to treat certain standards as unrevisable.

In fact, I think Worrall runs the thesis and the policy together, and attacks them both. I argue, however, that he does not manage to refute either of them. Thus, he establishes neither the unrevisability thesis nor the unrevisability policy. I will therefore frequently use variations of the somewhat inelegant disjunctive and conjunctive phrases "the thesis or policy" and "the thesis and policy", unless it is quite clear that it is one of them that is at issue at a particular point.

4.2 Worrall’s first objection: objectivity
Worrall formulates the first objection as follows:

If no principles of evaluation stay fixed, then there is no 'objective viewpoint' from which we can show that progress has occurred and we can say only that progress has occurred relative to the standards that we happen to accept now. However this may be dressed up, it is relativism. (Worrall 1988, 274)

Worrall claims that (a) unless some evaluative standards are absolutely fixed, there is no objective evaluative viewpoint; (b) that without such an objective viewpoint the distinction between a standard being accepted and its being normatively adequate (in some objective or absolute sense) collapses; and (c) unless such a distinction is upheld, rational evaluation is relativized to the standards of evaluation different agents happen to accept.

The objection is thus focused on the objectivity of evaluation and on the adequacy of evaluative standards. I think it is best viewed as making claims about how a proper idea of rational evaluation must be construed in order to avoid collapsing into an untenable rationality relativism.

My contention is that Worrall’s claims (c) and (b) are true, whereas (a) is false: there is no necessary connection between the objectivity of evaluations and treating evaluative standards as unrevisable (fixed). A proper non-
relativistic idea of rational evaluation does not presuppose that standards are unrevisable (or that we treat them as such).

Worrall is right when he says in (c) that if the distinction between merely accepted and really adequate standards is collapsed, the result is relativism. However, I find his claims about objectivity, and about the relation between objectivity and unrevisability (fixity) problematical.

By an "objective viewpoint" for evaluations, I assume that Worrall is thinking of a set of standards which are, themselves, objectively adequate or correct, in some sense — standards which are adequate independently of what particular agents believe about their adequacy. I agree that it is important to distinguish such a sense in which standards of rationality can be adequate (or inadequate), independently of whether we take them to be adequate or not — such a distinction is also important if we are to be able to make sense of the claim that we can be mistaken about such standards and that standards can be improved (see chapters 1 and 6).

The idea of objective evaluations is connected with the idea that the evaluative standards we use are themselves, in some sense, objectively adequate or inadequate. A standard is not adequate just because it happens to be my standard. If you will, a non-relativistic idea of rational evaluation requires some idea of the adequacy of standards, according to which that adequacy is not determined by what agents think of the standards.

However, making and respecting such a distinction — between an evaluative standard being accepted and its actually being adequate — does not require that we know for certain that any particular standards we accept actually are objectively adequate in this sense. Neither does it require that we treat them as unrevisable. Equipped with a distinction like this, we can claim that the standards we use, and thus the evaluations we make by appealing to them, are objectively adequate (just as we can make claims to objective truth), and we can do this without treating them as fixed. There is still the question of how we can argue rationally about whether standards are adequate or not, but the idea of objective, rational evaluation does not collapse into rationality relativism when it is divorced from the idea of unrevisable standards.
Therefore, the crucial step in the argument fails: contrary to (c), there is no necessary connection between the objectivity of evaluation and the unrevisability of standards. For us to be able to make the distinction discussed, it is not necessary that the evaluative standards we appeal to actually are objectively adequate, much less that we know with certainty that they are so or that we treat them as irrevocably fixed. In the terminology I introduced in chapter 1, we can have the concept of adequacy of standards, and distinguish between absolutely rational actions and standard-relatively rational actions. This we can do without knowing for certain which standards really are fully adequate and which actions really are absolutely rational.

The concept of adequacy of standards also allows us to distinguish between *mere change* and *progress* (or improvement) concerning standards: a change of standards constitutes progress if it leads to the adoption of standards which are more adequate than earlier ones. One can also, without invoking the idea of unrevisable standards, distinguish between mere change and *rational* change: one would judge that a change of standards is rational if it is permitted by standards that are taken to be adequate.\(^1\) Also, it should be noted that rational change and progress are not the same thing. It is one thing to evaluate a change of standards as rational (by appealing to some further standards). That such a change is also a case of progress means that the new standards are more adequate. A change may be rational but fail to result in improved adequacy. And a change may actually be an improvement although it is not rational.

But I believe that Worrall is getting at something important, both when he stresses the importance of objectivity and when he talks about the connection between objectivity and fixity. It is correct that rational inquiry and objective, rational evaluations require that in every evaluation there are standards which are at least *temporarily* fixed. It is also true that there must be some substantial continuity in the way we make evaluations. There has to be some *stability* among the things we appeal to in evaluating things. If things change too fast, we lose our bearings. This does not mean, however, that any standard must be treated as unrevisable (fixed) come what may.

\(^1\) Again, using the terminology I introduced in section 1.8, one can distinguish further between standard-relatively and absolutely rational change.
On the contrary, I cannot see how treating evaluative standards as unrevisable is of any help in trying to achieve objective evaluations. What matters, is whether the standards are adequate or not. "Fixing" them will not make them adequate. The risk is that, by deciding to treat a particular standard as unrevisable, we make sure that if it is not fully adequate we shall never be able to find this out and thus we will be stuck with an inadequate standard. Given that the possibility that a standard is inadequate can never be ruled out with absolute certainty (according to my argument in chapter 2), is it not wiser to refrain from "fixing" standards in this way?

My suggestion is that the best way to be objective about our evaluative standards is instead to keep them open to critical evaluation: to be open to criticism\(^2\) against them and to try to revise them if criticism shows them to be problematical (just as one does with beliefs or theories about the world).\(^3\)

This proposal amounts to the claim that the objectivity and rationality of evaluation is best served by treating all evaluative standards as open to critical evaluation and revision. Not only does Worrall's objection fail to establish the unrevisability thesis. His objectives are better served by adopting the revisability policy.

### 4.3 Worrall's second objection: independence

Here, however, we come to Worrall's second objection against the revisability thesis. Laudan argues against Worrall that the question of whether or not there are any fixed standards, or methods, is irrelevant for answering the relativist challenge. What matters is instead, according to Laudan, whether or not methods can be rationally evaluated, compared, and improved.\(^4\) Worrall now claims that rational evaluations and comparisons of

---

\(^2\) Remember that criticism is a form of evaluation.

\(^3\) Cp. Popper on scientific objectivity in Popper 1980, 44-46.

\(^4\) Actually, Laudan, in his reply to Worrall's review article, writes that what is required is that we be able to "justify" methods and to "show why certain methods are better than others" (Laudan 1989, 370). As far as I can see, Worrall's objection does not depend on interpreting the relevant sense of evaluation as being a matter of justification.

However, it should be added that Laudan nowhere gives an account of justification which shows how the trilemma of infinite regress, vicious circularity or nonrational commitment can be avoided.
different standards, and therefore rational revisions of standards, cannot be made unless there are some standards which are themselves beyond revision:

The question is whether our present point of view is right to say that our present methods are better than the methods of science of three centuries ago. And a positive to answer [sic] that question requires some principles considered as outside the historical process. (Worrall 1989, 381)

Worrall goes on to say that, if Laudan is to be able to claim that methods actually improve, rather than merely change, he must acknowledge that there are some fixed standards in terms of which such changes can be evaluated as rational improvements. Laudan, Worrall claims, to avoid a relativistic conclusion, must join "the 'fixed core' camp. To avoid joining the camp he must claim that even these judgments are grounded only within our present intellectual framework. But that position is classical historical relativism." (Worrall 1989, 381)

Here, Worrall repeats his claim that if with Laudan we hold all standards to be revisable, the distinction between merely accepted and objectively adequate standards is collapsed. That claim I have already criticized.

But this objection, as I interpret it, appeals in addition to the nature of non-question-begging evaluations. Worrall's claim is that standards of rational evaluation can be rationally evaluated and revised in a rational way, without begging the question, only if there is some core of standards which are themselves unrevisable.

As with the previous objection, I think Worrall points to something important and true about rational evaluations, but that this is not strong enough to support the claim that the revisability policy (or thesis) leads to relativism.

What Worrall says is very compressed, but I think the central point is that the evaluations in question must not be question-begging. In order to avoid begging the question in favour of particular standards that are involved in a

---

5 The second sentence in the quotation should begin: "And a positive answer to that question..." I interpret giving "a positive answer" as replying "yes".
revision, the evaluative standards used to evaluate a change of standards must be independent of, or neutral in relation to, the standards that are being evaluated. If one set of evaluative standards, S₂, is to be judged as better than another set, S₁, this requires appeal to some evaluative standard. Let us call this standard "Sₓ", and suppose that in the light of Sₓ, S₂ is judged to be better than S₁ (the change from S₁ to S₂ is judged to be a rational improvement). I think that Worrall’s point is that if one is wondering whether a change from S₁ to S₂ is rational or not, then one is not really helped by appealing to a standard that it would be rational to accept only given that one adopts S₂ instead of S₁. The standard which is used to evaluate competing sets of evaluative standards must be independent of or neutral with respect to the sets, or changes, involved, in the sense that it does not beg the question in favour of either of them. This can be summarized by saying that in order to be able to evaluate a change as rational, there must be some evaluative standard which is not itself at stake in this change. This demand, which I am here ascribing to Worrall, seems to be entirely reasonable.

But from this it does not follow that there is some standard which is presupposed in all evaluations of changes of evaluative standards (and therefore independent of all possible changes). What is required, is just that for rationally evaluating a change from one set of standards to another set as rational, the standard(s) appealed to in this evaluation must be neutral between the sets that are then being evaluated. This does not mean that there must be one standard (or set of standards) which is neutral with respect to every possible change of standards. From a statement of the form "In every rational evaluation there is a standard which is appealed to" one cannot derive a statement of the form "There is a standard which is appealed to in every rational evaluation" — just as from "For all husbands there is a woman to whom he is married" one cannot derive the statement "There is a woman who is married to all husbands".

If the objection is to constitute a valid argument, some further assumption must be added. Worrall does not supply such a premise, and without it, his objection does not constitute a refuting argument against the revisability thesis.
Before ending this section, let me try to give a different formulation of an argument that might be extracted from Worrall’s objection. Perhaps something like the following picture is behind the objection. First, let us form the set of all those standards which *can* be rationally revised. Then the question is, if this set can be thought to contain *all* our standards, or if there must be another set of standards which are rationally un revisable. The argument, then, would be as follows. We have the set of all revisable standards. But the rational revision of these requires that there is some standard (or set of standards) according to which these revisions are rational (or constitute improvements). And *this* standard cannot belong to the set of revisable standards, because if it did there would have to be a further standard which could be used to evaluate a revision of this standard too. That would mean that, contrary to the assumption, the standard in question would *not* be a standard for evaluating *all* revisions of standards. So, if any standards are revisable, then there must be some standard (or a set of standards) which is independent of any revisions and which is itself un revisable.

The assumption which is doing all the work in this argument is that there must be a single standard (or a single set of standards) which is responsible for evaluating *all* revisions of standards. But why should one make such an assumption? To evaluate a particular revision, say a change from standard $S_1$ to standard $S_2$, we need some standard which applies to these and which is believed to be adequate for the purpose. But why should we demand that this standard, in order to be adequate for evaluating a change from $S_1$ to $S_2$, must also be adequate for evaluating all possible changes between any two standards whatsoever? I do not believe that there is any such standard, but above all I do not think any such ”superstandard” is needed. Different standards will be relevant or applicable in different cases of revisions of standards. And without this assumption, this kind of argument for un revisability does not go through.
4.4 Unevaluability and relativism: turning the tables

Rationality relativism does not follow from the revisability thesis, or from the adoption of the (normative) revisability policy. At least the arguments I have discussed here do not show that it does. Now I shall try to turn the tables on the proponents of unrevisability. A central component in their case for unrevisability is the claim that some of our standards are such that they cannot themselves be evaluated. This *unevaluability* of standards would indeed have rationality relativism as a consequence. So I argue in this section.

In section 3.2 I argued that the limits of evaluation argument fails to show that some standards are unrevisable and incapable of being evaluated. I have argued in section 3.3 that, even if one accepts the principle of sufficient justification and the conclusion that necessarily there are limits to rational justification, it still does not follow that any standards of rational evaluation are forever beyond evaluation — or that these standards are unrevisable. At most, it supports the claim that at any given point there are some evaluative standards we have adopted or "committed" ourselves to in a nonrational way. It does not support the further claim that the standards are unevaluable and unrevisable, and that such a commitment is therefore *irrevocable*.

These claims about unevaluability of standards are crucial to the arguments in favour of unrevisability, but I think they are untenable if one wishes to avoid rationality relativism. It is precisely the thesis that standards of evaluation are unevaluable (and hence unrevisable) that leads to a form of rationality relativism, and it should therefore be rejected.⁶

As we have seen, some philosophers do claim that standards are unevaluable. Hauptli was led to this view by reflecting on the limits of evaluation. Others seem to be led to this claim as a result of reflecting on the sceptical challenge to rationalists — the challenge to justify the standards of rational evaluation. The limits of justification argument can be seen as a version of this attempt to deal with the sceptical challenge. In the previous chapter I quoted John Worrall’s claim that, in the face of the sceptic’s

---

⁶ As I stated in chapter 1, I regard it as a condition of adequacy for an account of rationality that it should not be relativistic. I limit myself to a discussion of non-relativistic accounts in this work, and will not argue directly against relativism.
challenge, "the only option is [...] the honest admission that ultimately we must stop arguing and 'dogmatically' assert certain basic principles of rationality." (Worrall 1989, 383) From the context, it is clear that Worrall holds the view that the standards we "dogmatically assert" in this way, must also be treated as unrevisable — as the "fixed core" of scientific methodology and rationality. Worrall himself sees this strategy as the only way of avoiding relativism. I believe, instead, that this way of answering the skeptic is bound to land you in the very relativism that Worrall (and other "fixity theorists") are anxious to avoid.

This (uncomfortable) position has the following components: (i) due to the limits of justification (or evaluation) we must ultimately accept some set of standards of rational evaluation in a nonrational way; and (ii) the standards we have thus accepted cannot themselves be rationally evaluated, since it is those standards that enable us to engage in rational evaluations in the first place.

The version of rationality relativism I have in mind consists in the following thesis: Different agents (or communities) adopt different sets of standards which lead them to different evaluations of beliefs, theories, etc., and it is impossible to rationally compare or argue about the adequacy of the different sets of standards that different agents adopt.7

The argument is very simple indeed. The relativistic thesis follows from (i) and (ii), given a further assumption: that different agents can and do accept different and incompatible standards of rational evaluation. As a matter of fact, this assumption is true. An example of two different and incompatible sets of standards might be those of a justificationist rationalist, who insists that everything we accept must be properly justified on the basis of rational arguments or experience, and those of a fideist who holds that a nonjustified commitment to some particular authoritative source is rationally permitted (or perhaps even required).

---

7 This formulation of relativism is compatible with an absolutist (non-relativistic) conception of truth — and even with the claim that standards of rational evaluation can be adequate, or correct, in an absolute (non-relative) sense. It means, however, that all claims about truth or adequacy can be evaluated only relative to different sets of standards of evaluation which cannot themselves be rationally criticized and compared.
Given that one has to adopt some standards in a nonrational way — so that prior to their adoption there can be no rational restraints on which ones to adopt — and that once adopted they cannot be rationally evaluated or revised, it would seem that which particular standards one adopts and continues to adhere to is not something that can be rationally discussed. And if different agents adopt different sets of standards and come to conflicting evaluations of something, then such conflicts could not be rationally resolved.

I might believe, of course, that my own standards of rational evaluation are the right ones, and that anybody who thinks differently is mistaken. This might even be true. But given (i) and (ii), I would be unable to either argue for such a claim or to examine it critically. Furthermore, to criticize someone else for appealing to a different set of standards would be pointless, since that would beg the question in favour of my own standards. And there is a perfect symmetry here: from the point of view of someone who had adopted different standards I would appear to be mistaken — and there would be no possibility of engaging in rational discussions of the conflicting positions. The theses of unevaluability (and unrevisability) would, if true, ensure that we are imprisoned forever by the particular standards we happen to accept — an acceptance which cannot itself be rationally discussed and evaluated. To adopt the unrevisability policy of "fixing" certain standards and treating them as immune to criticism and revision, unnecessarily risks putting us in exactly that position which the rationality relativist claims to be inescapable.

One philosopher who has argued that basic standards must be unevaluable, but that this does not lead to relativism, is A.J. Ayer. He discusses the rationality of induction, and the question of whether induction can be justified and defended against skeptical doubt. Precisely how an inductive standard should be formulated he does not go into, but says that in inductive reasoning "we make the assumption that there is a measure of uniformity in nature; or, roughly speaking, that the future will, in appropriate respects, resemble the past." (Ayer 1956, 72) With reference to Hume, Ayer says that

---

8 Unless, of course, we move from one position to another in a nonrational way (through leaps of faith).
this assumption, and hence the rationality of inductive reasoning, cannot be justified as true or even probable without arguing in a circle — that is, without relying on the assumption that is to be justified. Some form of empirical belief would be needed in a justification of induction. A belief that, for example, nature is uniform in certain respects, would itself have to be justified by some form of inductive reasoning. Similarly, if one argues that induction is justified because it works, one is again begging the question by presupposing induction: "It is tacitly to assume that the future can in this matter be relied on to resemble the past. No doubt this assumption is correct, but there can be no way of proving it without its being presupposed." (Ayer 1956, 74-75) So induction cannot be positively evaluated, that is, justified.

As I understand him, Ayer also thinks that it cannot be negatively evaluated, that is, rationally criticized, either. If we assume that positive and negative evaluation (justification and criticism) are jointly exhaustive, this would mean that induction cannot be rationally evaluated at all. His argument for the claim that it cannot be negatively evaluated seems to be that any rational criticism of induction would also presuppose induction:

This does not mean that the use of scientific method [induction] is irrational. It could be irrational only if there were a standard of rationality which it failed to meet; whereas in fact it goes to set the standard: arguments are judged to be rational or irrational by reference to it. (Ayer 1956, 75)

Here, Ayer seems to say that induction cannot be negatively evaluated, and that the unevaluability of induction does not have any disastrous consequences. But, as Bartley points out, Ayer is begging the question. He assumes not only that some standard of inductive reasoning is adequate and that inductive arguments are rational, but also that any rational criticism of induction must itself appeal to induction. But whether or not induction is rational, and whether or not there are any adequate competing standards, is precisely what is at issue here. As Bartley remarks:
Criticisms of putative standards of rationality have always questioned whether they were correct. Alternative conceptions of scientific method, such as Popper's which deny the existence of inductive procedure, let alone its legitimacy, do claim that there are standards of rationality which positions such as Ayer's fail to meet. (Bartley 1984, 99)

Without the question-begging assumption it does not follow that inductive reasoning is unevaluable (and thus unrevisable). But furthermore, even if it were the case that standards such as the one Ayer is discussing were unevaluable, the very same assumption is needed if Ayer's position is to avoid collapsing into rationality relativism. It is only by begging the question and assuming that induction is adequate, that Ayer can rule out the equally non-justified adoption of Popperian falsificationist standards, or indeed any standards. Ayer's version of the limits of justification argument does not show that there are unevaluable and unrevisable standards, and unless it is presupposed that some particular standards are adequate, the thesis of unevaluability does collapse into rationality relativism.

As I argued in chapter 3, the arguments from the limits of evaluation and of justification fail, and we are not forced to treat any standards as unevaluable or unrevisable. If the argument in this section is correct, this is fortunate, since given extremely plausible assumptions about the actual diversity of standards, unevaluability would have rationality relativism as a consequence. The rational option, accordingly, is to accept the revisability policy.

### 4.5 Conclusion

Worrall's objections, as I have tried to develop them here, raise important questions and express legitimate concerns about rationality in inquiry. But the objections do not amount to valid arguments with true premises showing that the revisability thesis entails rationality relativism. So they do not work as indirect arguments for the unrevisability thesis, either. Neither do they
show that the adoption of the revisability policy lands us in a relativistic position.

The revisability thesis says that what standards it is rational for an agent to accept is something that may change, and that such changes are possible concerning any standard. The revisability policy recommends that agents should be open to the possibility of rationally revising any standard they accept. Worrall is correct in claiming that such a view of rationality as revisable, would collapse into rationality relativism unless one insists on a distinction between a standard being accepted (by some agent) and a standard being adequate (objectively, whether or not anyone accepts it). But this distinction can be made (as I have done both in this chapter and in chapter 1) although one insists both that we can never know for certain if particular standards are adequate or not and that it is rational to treat all standards as revisable. There is no necessary connection between the concepts of adequacy and unrevisability, so neither the revisability thesis nor the revisability policy collapses into rationality relativism. At least the arguments examined here do not show that.

It is also true that the idea of making progress in revising standards requires some concept of the adequacy and inadequacy of standards. First, the idea that to replace one set of standards with a new set can be a case of progress rather than mere change, presupposes some idea of standards being better and worse, more or less adequate. And second, to rationally evaluate a change of standards requires an appeal to some standard which is not itself at stake in this change and which is taken to be adequate. But neither of these points are necessarily connected with the thesis that some standards are unrevisable or with a decision to treat some standards as immune to criticism and revision.

On the basis of the arguments I have examined and the arguments I have myself presented in chapters 3 and 4, I conclude two things. First, we are not forced to accept the unrevisability thesis. This does not show that its negation, the revisability thesis, is true, of course. Second, of the two competing normative policies, it is rational to adopt the revisability policy and refuse to treat any standards as in principle immune to criticism and revision.
In section 4.2 I proposed that Worrall's insistence that we should aim to be objective towards our standards, is best fulfilled by adopting the revisability policy. The decision to fix some standards and treat them as unrevisable indeed unnecessarily introduces an element of subjectivity which we, as long as we hold on to the unrevisability policy, cannot get rid of.

But how, then, does one rationally evaluate and revise standards? How does one rationally inquire into the adequacy of standards of rational inquiry? Can one say anything general about how one ought to go about this? This is the topic of the next chapter.
REVISABILITY AND BOOTSTRAP RATIONALITY

5.1 Introduction

In chapter 2 I argued that we are fallible with regard to all standards of rational inquiry (in the sense that there are no ultimate justifications which with certainty rule out the possibility of being mistaken about the adequacy of standards). In chapters 3 and 4 I criticized arguments for the claim that rationality presupposes that there are fixed or unrevisable standards, and argued that instead it is rational to treat all standards as open to criticism and the possibility of revision.

The general question to be considered in this chapter is: how can we account for the rational evaluation and change of standards if all such standards are held to be revisable? More specifically, I will try to answer the following two questions. (i) On the assumption that the revisability policy is to be adopted, so that no standard is to be set aside and treated as unrevisable, what does the process of evaluating and improving standards look like? (ii) Is it possible to formulate a theory about this process?

Briefly put, my answers to these questions will be as follows: (i) the process of evaluating and improving standards of rational inquiry is best seen as a "bootstrap" process; and (ii) a result of viewing it as a bootstrap process is that the process of evaluating and improving standards is not something one can have an informative and interesting theory about. If it is a matter of bootstrapping, then this process is open-ended, and cannot be legislated in advance.

I will argue that existing attempts to formulate bootstrap theories are inherently incomplete, in that any proposed resources for and constraints on rational evaluation and change they contain must themselves be regarded as provisional and revisable. However, these attempts to formulate theories can nevertheless give valuable contributions to our understanding of bootstrap operations and of how evaluative standards,
methodological rules and cognitive goals can be rationally evaluated and revised.

A significant part of this chapter will be devoted to an examination of two attempts to formulate a so called "bootstrap theory" of rationality in inquiry.

5.2 Rational standard change as a bootstrap process

At every stage of inquiry in any field, there will be a diverse set of standards of different kinds. Some of these may themselves be inapplicable to other standards, while others may instead be metastandards applicable only to standards. There will also be methodological rules, logical inference rules, cognitive goals, and a vast number of beliefs both about the subject matter being studied and about the inquiry into that subject matter.

If one wants to evaluate one of the current standards, or a proposed change of the current set of standards, then this evaluation has to be made in the light of some of the standards that one already possesses. This is so, since every evaluation presupposes some standard, and since a genuine normative evaluation presupposes that the standards we appeal to actually are ones that we accept.

What should a general characterization of this process of standard evaluation and change look like?

One alternative would be to propose that there is a fixed core of standards which themselves are unrevisable, but which can be used to evaluate changes of other standards. This would be what I will call a "fixed core" view of standard change. Ideally, such a view should be concretized into a theory which would make these core standards explicit and give detailed directions for how they are to be brought to bear on the evaluation of changes of other, more peripheral standards.¹

The idea behind a core theory would be that the core standards are themselves unrevisable, and thus constitute a fixed point of evaluation from which judgments about the rationality of changes may be made.

¹ Two contemporary philosophers who hold the fixed core view, claiming that rationality necessarily requires that some of our standards are unrevisable, are Worrall and Wedgwood. See Worrall 1988 and 1989, and Wedgwood 1999. Neither of them has specified this core in any detail, though.
However, I argued in chapters 3 and 4 that there are no good reasons which force us to assume that any of our standards of rational inquiry are thus unrevisable. Instead, I further argued, the rational policy is to treat all our standards as open to critical evaluation and to the possibility of future revision.

A consequence of this revisability policy is that a general account of the rationality of changes of standards of rational inquiry should not include any assumption about unrevisability. Thus it should not contain a distinction between a core of unrevisable standards and a periphery of revisable ones. According to the revisability policy, we should treat all standards as criticizable and revisable, even the standards we use for evaluating standards. Fixed core theories should therefore be rejected along with the view of the process of rational standard change as dependent upon fixed standards.

An alternative is to view the evaluation and change of standards as a bootstrap process.2 That is what I propose to do. But what does it mean to say that it is a bootstrap process? What is that?

"To pull oneself up by the bootstraps" is a metaphorical expression. Taken in the literal sense it is, of course, impossible to pull oneself up just by the bootstraps. Joseph Agassi offers the following elaboration of the bootstrap metaphor, contrasting it with the picture of someone climbing a mountain or a ladder:

The mountain climber (and the ladder climber) makes sure to the best of his ability that a step he is going to take will put his forward foot on solid grounds. Of course, the mountain climber can never make absolutely sure that his next step will be safe until he has put it behind him. Yet, his success in making any step rests on the stability of the place on which he has rested his foot in the previous step. In a bootstrap operation the opposite seems to be the case: the support cannot sustain one well enough, yet one uses it just long enough to make the next one. Literally, of course, the image is exaggerated to the point of impossibility — we cannot pull ourselves by our bootstraps.

---

2 The bootstrap view of rational change of standards discussed in this chapter is not, as far as I can tell, related to Glymour's bootstrap theory of confirmation (see Glymour 1980) or to the statistical bootstrap method (see Efron and Tibshirani 1993).
alone — yet the allusion is to ever so many situations which are notoriously precarious yet which we voluntarily enter in the hope to exit very quickly from, and into a better position than ever before. An example may be walking fast in a swamp or a speculation in a bull market by a penniless customer who, however, has enough credit to speculate; the customer resells at a higher rate before he has to pay. (Agassi 1975a, 169-70)

Other contrasts to the metaphor of pulling oneself up by the bootstraps, might be the metaphor of climbing a pole which is firmly planted in the ground below you, or perhaps climbing a rope which is securely fixed at some point above you. The bootstrap metaphor, as I understand it, is supposed to convey a negative point: that improvement, and rational judgments about improvement, do not require the existence of any unchanging, fixed evaluative standpoint.

The bootstrap idea can be formulated as follows: we use a subset of our current evaluative standards for evaluating certain other standards (or, possibly, to criticize a self-applicable standard by applying it to itself), for comparing competing alternative standards, and for judging whether proposed changes of standards constitute improvements. This evaluation procedure may in principle come to be used on any particular standard among those we adhere to. Specifically, those standards which we use at one point to evaluate other standards, may themselves later on come under scrutiny and be revised, possibly in the light of new standards.

In slightly different terms, the bootstrap view says that we use a subset of our standards to evaluate a second subset and to change our total set of standards; this results in a new total set, a subset of which is used for further evaluation and improvement, and so on.

According to the bootstrap view we use some of our present standards to evaluate and improve our own standards. What our evaluations amount to and which improvements are judged to be either required or possible, depends on what standards we actually possess at a particular time, as well as on what other resources are available to us at that point (including, e.g., our awareness of alternatives or of specific arguments for and against particular standards). Thus, in the bootstrap process the very standards we use for evaluating standards and change of standards, may themselves change.
This bootstrap view of standard change has an important negative characteristic: it does not require that a series of changes of standards must be made in the light of a single set of (meta-) standards; it allows that a series of changes may be rational without the presence of any constant core of standards. The bootstrap element consists precisely in the idea that we may at one point rationally use standards or beliefs which at later stages of the process come to be discarded.

A further consequence concerns the prospects for formulating a theory about the bootstrap process. Bootstrapping is not the sort of thing you can have an interesting theory about. I shall explain why this is so.

I will not attempt to explicate the meaning of "theory" or of "interesting theory". However, I propose that in order to count as a genuine, interesting theory about the bootstrap process, that theory should satisfy two desiderata. For want of better words, I refer to these desiderata under the headlines "completeness" and "restrictions". My claim is that the bootstrap view itself entails that no theory can satisfy them.

The *completeness* desideratum means that a theory should be able to specify the kinds of considerations that can be used in making judgments about rational changes of standards. The theory should say something quite general but still specific about what kinds of evaluations and changes are allowed as rational bootstrap operations. It should be (potentially) complete in the sense of (potentially) exhaustive.

The *restrictions* desideratum means that a theory about bootstrap processes, in order to say anything informative, should also rule out, in a principled way, certain kinds of changes as irrational.

The bootstrap view discussed above cannot be worked out into a completely general theory which satisfies these desiderata. The reason is that part of the bootstrap view is that any part of our present set of standards may at a later point come up for revision. Similarly, of course, it is also taken to be a possibility that new standards may come to be proposed and added to our set of standards.

This means, first, that a theory about the bootstrap process could not even aspire to be complete, since according to the bootstrap view itself it is possible to add new standards and thus new kinds of considerations concerning the rationality of changes. Even if it were possible to lay
down a list of kinds of changes that are allowed as rational, this list would have to be regarded as open-ended.

Second, according to the bootstrap view any of the standards now believed to be adequate may, in principle, come to be revised further on. Any particular standard is open to becoming either completely rejected, replaced, or in some way reformulated as a result of further inquiry. This means that considerations and kinds of changes that had previously been allowed as rational, may come to be disallowed. That is, the "list" would be open also in the sense that any item on it may come to be reformulated or simply deleted.

This second remark also applies to the question of the restrictions a theory would make — what changes it rules out as irrational. What these restrictions are, depends on what standards are deemed appropriate for evaluating changes of standards. But since the bootstrap view itself says that any such standard is to be treated as in principle revisable, then any particular standard which at present is taken as a restriction on rational bootstrap operations may, in principle, itself come to be dropped further on.

As far as I know, it was Agassi who introduced the idea of bootstrap rationality in Agassi 1974. In that paper he does use the term "bootstrap theory", but to my knowledge he has not himself attempted any detailed and comprehensive formulation of such a theory.

Other theorists have, however, tried to put some meat on the admittedly meagre bones of the bootstrap view. In later sections of this chapter, I shall discuss two such accounts of bootstrapping: Larry Laudan’s "reticulated model of scientific rationality" and Larry Briskman’s "bootstrap theory of rationality". Given the considerations above, neither of these will qualify as a real theory of the bootstrap process — since there can be no such theory.

Nevertheless, accounts such as Laudan’s and Briskman’s can perhaps contribute to our understanding of bootstrapping. They may indicate to us ways in which standards of rational inquiry can, given certain assumptions about rationality, be evaluated, compared and improved. Although they are not complete theories, they provide general frameworks within which different kinds of rational evaluations and changes of standards can be explained.
Before I turn to a discussion of Laudan and Briskman, I shall briefly present the positions of two other philosophers — positions that indicate at least the role of a bootstrap view of rationality, but which do not themselves contain any statement of such a view.

5.3 Bartley and Siegel
In *The Retreat to Commitment* (1984), Bartley discusses the problem of the limits of rationality. He argues that this problem arises for any position according to which rationality is a matter of justification. More specifically, the attempt to justify rationality itself, and the standards of rational justification, threatens to open up an infinite regress of justifications. The result, according to Bartley, is that as long as rationality is seen as requiring justification, a rationalist must admit that rationality and standards of rationality is something to which one must ultimately commit oneself irrationally (or nonrationally).

This enables irrationalists (Bartley’s example is of forms of fideist irrationalism in modern Protestantism) to use the so called "*tu quoque*" ("you too") argument against rationalists: whenever a rationalist criticizes an irrationalist for believing things or acting in an irrational way, the irrationalist can answer "*tu quoque*". Due to the limits of rational justification and the consequent need for an irrational ultimate commitment to rationality and to standards, the rationalist is, so to speak, in the same boat as the irrationalist.

Bartley’s proposed solution is to separate criticism from justification. He develops a theory of rationality, pancritical rationalism, according to which rationality is a matter only of criticism. A pancritical rationalist is not forced to treat anything as beyond criticism — hence there are no limits to rationality and no need for any nonrational commitments. The *tu quoque* cannot be used against this position.

Bartley presents the pancritical rationalist as follows: he is someone "who is willing to entertain any position and holds *all* his positions, including his most fundamental standards, goals, and decisions, and his basic philosophical position itself, open to criticism" (Bartley 1984, 118).

In a comment on the status of the evaluative standards of criticism a pancritical rationalist would use, Bartley says:
we must be willing to reopen to examination and further criticism and possible rejection all the critical arguments and critical institutions we have accepted. [...] Such a willingness to hold open to revision in principle even those notions that we believe most surely to be true is part of the spirit of pancritical rationalism. (Bartley 1984, 126)

Bartley is of course aware that a critical argument always presupposes something, that one must temporarily accept some things in order to be able to criticize others. He is also aware that rational criticism requires appeal to some standards of rational evaluation. And it is part and parcel of pancritical rationalism that all evaluative standards are themselves to be held open to critical evaluation and revision.

According to Bartley, rational inquiry and discussion is a matter of criticizing statements or positions. At every stage this criticism uses some standards of critical evaluation. These standards are themselves criticizable and revisable. At some later stage of the discussion, the standards used at earlier stages may in turn be criticized and possibly rejected.

Thus, Bartley’s formulation of pancritical rationalism contains some of the elements of a bootstrap view of rationality. However, he does not say much more than this about how standards of rational evaluation are evaluated and revised. An elaboration of the bootstrap view would aim to give a fuller account of the "dynamics" of standard change.

The need for a bootstrap view of rationality can also arise in justificationist accounts of rationality. As an example of a justificationist philosopher whose position implies such a need, I shall take Harvey Siegel.

Siegel’s *Relativism Refuted* (1987) mainly consists of criticisms of different forms of epistemological relativism. In the last chapter, however, he discusses what an adequate non-relativistic epistemology (which Siegel calls "absolutism") should be like. One of his aims is to make clear that relativism and what he calls "vulgar absolutism" are not

---

3 Critical rationalism in general — of which Bartley’s pancritical rationalism is one version — implies some sort of bootstrap view of rationality.
4 Other examples of this might be Israel Scheffler (see Scheffler 1973, 80), and Jürgen Habermas (this might seem surprising in his case, but see the discussion of Habermas in chapter 2).
the only alternatives. A vulgar absolutism would claim that there is a "unique privileged framework" (Siegel 1987, 161) of standards, which are both infallible and incorrigible. Such a "vulgar" theory Siegel finds untenable, but he claims that an absolutist (i.e., non-relativistic) epistemology without these "undesirable features" (Siegel 1987, 164) is possible:

Absolutism requires that it be possible that claims be evaluated in a non-question-begging way, and that objective comparison of rival claims be possible. It does not require that knowledge be certain, or indubitable, or rest on considerations to which we have privileged access. For the criteria which afford objective, non-question-begging evaluation may themselves be criticized, compared with alternatives, and improved. (Siegel 1987, 165)

Siegel does not, however, say very much more about what the features of an absolutist epistemology should be, and he does not discuss how its standards "may themselves be criticized, compared with alternatives, and improved" (Siegel 1987, 165). What Siegel claims is that any of our present (and future) standards can be held open to future criticism and revision, although at each stage of inquiry we must appeal to some standards or criteria. As far as I can see, this process of criticizing and evaluating criteria of evaluation would be a bootstrap process (on pain of self-contradiction on Siegel’s part). But Siegel himself does not provide any account of that kind of process.

In the remainder of this chapter I will discuss two attempts to give such an account, that is, attempts to explain how evaluative standards and cognitive goals may be rationally evaluated and changed in the absence of absolutely fixed standards.

5.4 Laudan's reticulated model of scientific rationality
In earlier chapters, I have referred to Larry Laudan’s claims that all standards (methodological rules) are fallible and revisable. Laudan has also developed something he calls the "reticulated model of scientific rationality", which not only is supposed to be compatible with the fallibility and revisability of standards of rational evaluation, but also
explain how such standards can be rationally evaluated and changed — without assuming that there are any fixed scientific standards. This model, which is presented in *Science and Values* (1984), contains some of the elements of the bootstrap view of rationality.

In *Beyond Positivism and Relativism* (1996), he goes on to develop his ideas about the evaluation of standards and methods, and proposes a more specific meta-methodology for testing methodological rules and evaluative standards. I shall treat the "reticulated model" and these specific meta-methodological proposals separately.

**The reticulated model**

Laudan himself uses the bootstrap metaphor to characterize the way in which scientific rationality, according to him, develops. As I explained earlier, bootstrapping is not the sort of thing you can have a real theory about. Here I shall instead examine to what extent Laudan provides resources for understanding ways in which rational bootstrap processes *can* occur. The bootstrap character of his view is clearest in the case of his reticulated model.

Laudan distinguishes between three different levels in science: factual claims, methodological rules (which include what I call "evaluative standards" as well as "methodological rules"), and the axiological level of cognitive aims (or goals). In *Science and Values* he expounds his own view by contrasting it with a somewhat simpler one which he calls "the hierarchical view". I shall follow his example.

The central idea of the hierarchical view is that methodological rules are regarded as means to scientific goals, and that such rules can be evaluated by examining how effective they are as means to given goals. The view is hierarchical, in the sense that it is one-directional: factual claims are evaluated by appealing to methodological rules, which are in turn evaluated in relation to goals. Goals are, according to this view, at the top of the "justificatory ladder", and are themselves seen as beyond rational evaluation.

Laudan himself accepts the goal-directed, or instrumental, view of methodological rules. Methodological rules are, Laudan agrees, to be seen as means to cognitive goals.

What he objects to is precisely the hierarchical or one-directional nature of the hierarchical view. For one thing, he argues that factual
claims (both about the world and about ourselves as inquiring agents) can be relevant to and constrain our choice of methods. More importantly, he claims that scientific goals can also be rationally evaluated, and that a theory of scientific rationality which does not allow for this is seriously defective.

Laudan describes his own model in the following, rather loose and casual way: "there is a complex process of mutual adjustment and mutual justification going on among all three levels of scientific commitment. Justification flows upward as well as downward in the hierarchy, linking aims, methods, and factual claims. No longer should we regard any one of these levels as privileged or primary or more fundamental than the others." (Laudan 1984, 62-63) In a figure, entitled "The Triadic Network of Justification", Laudan gives a schematic account of these interrelations and mutual dependencies. Methods justify theories and exhibit the realizability of aims. Aims justify methods. Theories constrain methods, and theories and aims must harmonize with each other.

In Laudan's model, no element needs to be taken as absolutely fixed. What he does, is to separate out different kinds of elements in scientific inquiries — factual beliefs, methods and goals. These elements stand in various relations to each other, and can, according to Laudan, be used for evaluating each other. The idea is that at particular stages of inquiry some particular elements — say, factual beliefs and goals — can be treated as temporarily fixed and used for the purpose of evaluating and judging changes of other elements — e.g., evaluative standards. At a later stage, new standards may instead be used to evaluate factual beliefs or goals, and so on.

The reticulated process of change thus seems to be a kind of bootstrap process — a process of piecemeal, gradual change where evaluations are made in terms of elements that may themselves come to be evaluated and changed at later stages. That much was contained in the rather thin characterization of bootstrapping that I suggested earlier in this chapter. What does Laudan’s model contribute when it comes to giving a more substantive account of bootstrapping?

For one thing, the reticulated model is supposed to deal with the evaluation and change not only of standards and rules, but also of cognitive goals. Also, my earlier characterization was formulated in terms of standards that are used as temporarily fixed for the purpose of
evaluating changes of other standards, whereas Laudan’s model explicitly incorporates factual beliefs and goals and the roles that these play in evaluations. Thus, by bringing these items in, and discussing how they are related to and can be used for the purpose of evaluating changes of standards, the reticulated model may be seen as an attempt to put some meat on the bones of the bootstrap view. So far, so good.

As far as I can see, Laudan suggests three things: (1) factual beliefs (about the aspect of the world we are investigating or about ourselves as inquiring agents) can sometimes be used to criticize ("constrain") existing methods. (2) It is possible to criticize goals, and to change goals rationally on the basis of criticism. (3) Methods (evaluative standards and methodological rules) are used in order to enable us to achieve goals, and can therefore be evaluated by seeing whether or not they actually function as means in our goal-pursuit.

Concerning the first point (1), how factual beliefs can function in the evaluation of methods, Laudan does not say very much. One of his examples is that beliefs about the world we live in, to the effect that "nature does not offer information to us in a random or statistically representative way" (Laudan 1984, 38), has rationally motivated the development of random sampling techniques in order to obtain representative evidence. Another example is how the empirical discoveries of the placebo effect and the effect of therapeutic expectations motivated the introduction of single- and double-blind testing of pharmacological drugs. The idea is that in such cases the application of current standards and methods leads us to the discovery that there are possible sources of error in our inquiries as hitherto conducted, and that we should revise standards so as to avoid these.

There are of course factual beliefs of different kinds. The ones Laudan mentions here as examples of beliefs that "constrain" methods (e.g., beliefs about the placebo effect) are beliefs that are somehow about the objects or the domain that is being studied. But we also have beliefs about ourselves as inquiring agents (our capabilities and limitations)\(^5\) and about the inquiries that we are engaged in. The latter kind of belief, empirical information about the history of inquiries, plays an important role in Laudan’s account of how methods (standards and rules) are evaluated.

\(^5\) For an interesting discussion in which it is argued that the concept of what is observable evolves along with science, see Shapere 1982.
The idea (2) that standards are means to goals of inquiry is important, I think. It introduces a suggestion about what the adequacy of standards and rules consists in; by saying what standards are supposed to be good for, it gives us an idea at least of how standards might be evaluated, compared and rationally changed. This I will refer to as "the instrumental conception of adequacy". It is, of course, not an original idea. As a claim about what the adequacy, or goodness, of standards and methods consists in (at least partly), it is not at all controversial and would be accepted by most methodologists and epistemologists. What is more original and controversial, are the meta-methodological proposals Laudan makes in Laudan 1996, where he actually suggests how to go about testing and evaluating the adequacy of rival standards and methodological rules. I shall discuss his idea about how to evaluate standards and methods in terms of goals in the context of that meta-methodological work in a later section.

I will discuss Laudan’s account of goal evaluation, (3), in the next subsection. Before that, I shall make some general comments on the reticulated model and the way in which it relates to the bootstrap view of rationality.

First of all, things are a bit more complicated than Laudan’s picture of reticulation suggests. In every kind of evaluation some evaluative standards have to be used. This is the case also, for example, when methods are evaluated in relation to goals, and when goals are criticized. Thus, the kind of bootstrapping, or reticulation, Laudan envisages requires that there is a multiplicity of different kinds of standards (for different purposes) that are applicable in the different sorts of evaluations included in his reticulated model. If Laudan’s model is to work, there has to be reticulation that takes place at the level of standards.

Concerning the evaluation of standards in relation to cognitive goals, for example, one might propose different evaluative meta-standards. One might require that, for a standard or rule to be rationally acceptable, one must be able to show that it is indeed a means to the goal in question. This is, as will be seen later on, what Laudan himself proposes. An alternative proposal, is to allow that a rule or standard is rationally acceptable as long as it has not been shown to frustrate the attempt to achieve the goal (something like this is suggested in Briskman 1983).
The same thing applies to the evaluation of goals. When one criticizes a goal, one always does so in the light of some evaluative standard. Laudan himself proposes that if one is to be rational in adopting and pursuing a particular goal, one must have positive reasons for thinking that this goal is indeed achievable — that is, that one must be able to justify the claim that the goal can be realized. This is itself a kind of standard for evaluating goals, a standard which can also be further specified in different ways (that is, the justification requirement can be specified in various ways, which leads to different standards of goal evaluation). Alternatively, one might instead propose non-justificationist standards of goal evaluation which allow that a goal is rational as long as it has not been effectively criticized and shown to be unachievable (see Briskman 1983).

Laudan is right, I think, to emphasize both that factual beliefs and goals may be used in evaluating standards and that goals can themselves be evaluated. However, it is misleading to say that methods (if this is understood in a wide sense so that it includes all evaluative standards and methodological rules) can be rationally evaluated and changed if goals and factual beliefs are treated as temporarily fixed. This is misleading because some element from the category of methods, some evaluative standard or other, is involved also in the evaluation of methods (standards and rules). To evaluate our standards, we need to use some standard and treat it as temporarily fixed. He does not deal explicitly with the fact that all evaluations require appeal to some evaluative standard. So even if Laudan is right in saying that factual beliefs and goals play important roles in such evaluations, the picture he paints is too simple.

Nevertheless, Laudan’s accounts of goal and method evaluation does say something about how goals and methods can, in certain circumstances and given certain assumptions, be rationally evaluated and changed. I shall therefore devote the next two sections to a discussion of these accounts.

Laudan on goal evaluation
It is interesting to note, that Laudan’s account of goal evaluation is an account of how proposed goals can be criticized. He does not seem to think that goals are susceptible of justification (as he does concerning methods and factual beliefs). He writes that there are "two general modes of criticizing a proposed cognitive goal or set of goals (apart from
charging it with inconsistency). I will show that one may argue against a
goal on the grounds (i) that it is utopian or unrealizable or (ii) that it fails
to accord with the values implicit in the communal practices and
judgments we endorse.” (Laudan 1984, 50) He adds that these do not
"exhaust the resources of axiological critique” (Laudan 1984, 50) but that
they are the most important ones.

He distinguishes between three ways in which a goal can be (criticized
for being) utopian. (a) ”Demonstrable utopianism”: ”We can sometimes
show that a certain cognitive goal cannot possibly be achieved, given our
understanding of logic or the laws of nature.” (Laudan 1984, 52) One
example he uses to illustrate this, is that of someone who has the goal of
travelling at speeds exceeding the speed of light. As an example in the
sphere of cognitive goals, Laudan takes the abandonment, during the 19th
century, of the goal of infallible knowledge (at least in the case of
universal theories).

It is not quite clear what this kind of criticism, and the reference to
”our understanding of logic or the laws of nature” amount to. It is not
supposed to be a matter of showing that goals are inconsistent. But then it
is not clear to me why he refers to logic. And the reference to laws of
nature seems curious, too. Consider the example of an Aristotelian
scientist who believes that physical changes are controlled by essences
inhering in things, and that physical processes are teleological. Such an
Aristotelian might argue that it is impossible to give purely mechanical
explanations of such changes (since such a goal is incompatible with his
beliefs about the laws of nature). It can hardly be this kind of argument
Laudan is thinking of.

It actually seems to be something weaker than demonstration that
Laudan has in mind. In the example he himself gives, the rejection of the
goal of infallible knowledge, the reason for rejecting infallibilism seems
not to have been that the goal itself was logically flawed or impossible in
the light of natural laws, but instead that no methods that had ever been
proposed (such as Cartesian or Aristotelian intuition of self-evident
principles, or eliminative or enumerative induction) could plausibly be
claimed to lead to infallible knowledge.6

6 Maybe Gödel did demonstrate that the goal of providing finitistic proofs of the
consistency of mathematical theories is impossible to achieve?
(b) "Semantic utopianism": this is when proposed goals cannot at all be adequately characterized, and are too imprecise or ambiguous to play any role in actually guiding inquiry. He tentatively suggests that such goals as simplicity and elegance might be semantically utopian.

It should be noted that the rational outcome of a criticism to the effect that an alleged goal is semantically utopian, need not be a wholesale rejection of it. Another rational response might be to revise the characterization of the goal, so as to remove ambiguities or make it more precise, so that it becomes capable of playing a role in guiding inquiry. Popper, for example, regarded Tarski’s work on truth as such a "rehabilitation" of the concept of truth (see Popper 1972, 320-24). This is also what Popper himself — tentatively — proposed to do concerning the concept of simplicity (Popper 1980, ch. VII).

Concerning the examples Laudan mentions here, simplicity and elegance, one can also wonder whether these should really be regarded as goals, or if they rather belong to the category of evaluative standards. It is not uncommon to view simplicity, for example, as an evaluative standard in terms of which rival theories may be compared. Are simplicity and elegance goals or standards, or are they somehow ambiguous? I think that there is a simple explanation of this: that often there is no sharp distinction between evaluative standards and goals. A reason for this is that some standards may be partially constitutive of goals, since they function as specifications of these goals. For example, standards of adequate explanation serve in part to specify the goal of having true explanatory theories, while at the same time they are used to evaluate competing theories. Similarly, if part of one’s goal is simplicity — to find simple theories — then it is natural to appeal to some standard of simplicity in evaluating competing theories. The same thing can happen if goals are specified in epistemic terms (e.g., the goal of having empirically justified theories, or demonstrated truths). There is nothing strange in this, but I think that it constitutes a difficulty for Laudan’s meta-methodological theory. The reason is that his account of the empirical evaluation of methods, in which one tests claims about contingent means-ends relations, requires that standards and goals be kept clearly separate.

(c) "Epistemic utopianism": A goal is epistemically utopian if there is no, explicit or implicit, ”criterion for determining when the value is
present or satisfied and when it is not.” (Laudan 1984, 53) Laudan devotes the final chapter of *Science and Values* to arguing that truth is epistemically utopian in this sense as a goal of scientific inquiry. One should note, however, that this is a very strong, and controversial requirement for the appropriateness of goals. A realist may well admit that there is no "criterion for determining" if a theory is true, or closer to the truth than some other theory, but that the search for goals such as truth can be rational in the absence of such criteria.7

Finally, there is the critical strategy (ii) of arguing that proposed goals fail to accord with the values implicit in the actual actions and judgments of the scientists concerned. As Laudan observes, there are great difficulties in stating exactly what it is for a goal to be implicit, and in identifying such implicit goals. However, he says, "there are often tensions between an agent's avowed or explicit goals and the goals that seem to inform his actions. Because there are, it becomes possible to criticize an agent's explicit goals by pointing out how contrary they are to the goals that evidently undergird his actions and practical judgments.” (Laudan 1984, 54)

Now, Laudan is not arguing that we should always adjust our explicit axiology so that it conforms to our implicit goals — he does not claim that practice is always primary in relation to theory. What is revealed by these kinds of critical arguments, is that there is an inconsistency or tension between the explicit and implicit axiologies. That shows that something needs to be revised, but it need not tell us what to revise: "Whenever a case can be made that a group of scientists is not practicing what it preaches, there are prima facie grounds for a change of either explicit or implicit values. The change may come, of course, in either area, or in both.” (Laudan 1984, 55)

Laudan does not say so explicitly, but the same thing seems to be the case with the critical strategies (a) and (c) above, too. Remember that for Laudan methods can be criticized for being inadequate as means to our goals. To argue that a goal is "utopian" in any of these senses, amounts to claiming that there is no known method (standard) which allows us to make rational judgments about whether we achieve, or are approaching, the goal in question. But then this can be taken *either* as a criticism of the

---

7 See, e.g., Popper 1962, 369-74, for a clear statement of this view.
goal, or as a criticism which shows that we need to develop better methods and standards.

On the whole, Laudan’s account of how to rationally criticize and change proposed goals of inquiry is not all that exciting. It consists in appeals to what is taken to be general features of goal-directed rationality — that it is problematic to aim at what one believes to be patently impossible, and so on. It is, one could say, an instrumental view of the rationality of goals. Given that one understands rationality in inquiry as goal-directed, Laudan does indeed have something to offer: he gives examples of ways in which goals can indeed, under certain circumstances, be criticized and revised.

But it must also be noted that Laudan’s suggestions about how to evaluate goals presuppose some standards of goal evaluation. The kind of criticism which says that a goal is epistemically utopian, for example, presupposes that in order for a goal to be rational there must exist some criterion for determining when it is realized and when it is not. This is a rather strong standard, which can be discussed, and perhaps rejected. For instance, no realist who is also a fallibilist would accept it as an adequate standard of goal evaluation, I think. Laudan’s reticulated model would have to deal explicitly also with the question of how to rationally evaluate and change standards of this kind, in order to say something interesting about bootstrap rationality. It does not do this.

Laudan on empirical testing of methodological rules

In Laudan 1996 he develops his idea of how methods can be tested by examining how they contribute to the achievement of cognitive goals. While the reticulated model described above is highly general and contains some of the elements of the bootstrap view of rationality, the account of how to test methodologies to be discussed in this section is rather a contribution to a different kind of theoretical enterprise: it is a meta-methodology which aims to provide concrete guidelines on how to test and evaluate proposed standards of rational scientific inquiry.

Laudan observes that methodological rules usually ”have the form of commands” and that they often have the categorical form ”‘you (or one) ought to do y’” (Laudan 1996, 134). He believes that this way of formulating them is misleading, and that they are more properly formulated as ”hypothetical imperatives whose antecedent is a statement
about aims or goals, and whose consequent is the elliptical expression of the mandated action.” (Laudan 1996, 132) So methodological norms rather have the following form: ”If your (one's) central cognitive goal is $x$, then you (one) ought to do $y$.” This he refers to as a rule of ”type (1)” (Laudan 1996, 134).8

Such a rule is normative, and has the form of a conditional imperative. Such rules are generally taken to be neither true nor false. However, it asserts a connection between means (actions of a certain type) and goals, and thus in some sense depends on statements which can be true or false. Therefore, Laudan claims, such rules are ”warranted” only if the statement ”Doing $y$ is more likely than its alternatives to produce $x$, ’ is true or warranted.” (Laudan 1996, 134) That, Laudan writes, is ”a conditional declarative statement, asserting a contingent relationship between two presumably 'observable' properties, namely, 'doing $y$' and 'realizing $x$'. Specifically, [it] has the familiar form of a statistical law.” (Laudan 1996, 134). Such a statement can be empirically tested, Laudan claims, by examining the results of past applications of methodological rules.9

If Laudan is right (or, to the extent that he is right) this explains how evaluative standards can be rationally evaluated and revised: standards are evaluated (in a comparative manner) by examining their track records — by looking at the proportion of cases in the past in which using a certain standard has led to the realization of a given goal. A revision is rational if the new standard which comes to be adopted has led to the realization of the goal in a higher proportion of cases, than did the standard it replaced.

Laudan’s idea of empirical tests of methodological rules has certain problems and limitations, I think.

---

8 This formulation is problematic. Laudan himself emphasizes that one and the same rule can be a means to more than one goal (Laudan 1984, 44-46). But this requires that the rule can be formulated independently of particular goals. Therefore, the statement of the goal should not, I think, be included in the formulation of the rule or standard. Instead, one might follow Nola and Sankey in distinguishing between methodological rules, methodological principles, and values (goals). A rule tells one to act in a certain way (for example, "reject unfalsifiable theories"), whereas a methodological principle is a hypothetical imperative which links a rule and a value: "if one wishes to realise value $v_i$ then one ought to follow rule $r_j$" (Nola and Sankey 2000, 9).

9 Since the ”warrant” of methodological rules depends on statements which have truth values, Laudan is himself prepared to speak of methodological rules as being true or false.
(a) One limitation, which Laudan himself emphasizes, is that goals often "underdetermine" the selection of standards (Laudan 1984, 35-37). The available evidence may be insufficient to decide between standards which compete as means in relation to a given goal.

(b) Laudan also admits that this kind of empirical testing of standards does not exhaust the possibilities of standard evaluation. In his reply to Doppelt (Doppelt 1990) he makes clear that methodological rules can also be evaluated by non-empirical (conceptual) arguments (Laudan 1996, 160-61), and also writes that there might be "ends/means connections which are, in effect, analytic and whose truth or falsity can be established by conceptual analysis." (Laudan 1996, 261 n. 20) So, even if the kind of empirical testing Laudan discusses is one possible way of evaluating methods, it is not altogether clear how central it really is in relation to other kinds of evaluation.

(c) Laudan’s model of empirical standard evaluation seems to presuppose that it is rather unproblematic to determine both that a certain kind of action has been performed, and that this action has lead to a certain result (realizing a certain goal or failing to do so). It should be noted, however, that this does not work for such "ambitious" goals as, e.g., truth or verisimilitude. Conversely, since goals are according to Laudan evaluated by examining whether there exist any viable methods for achieving them (and criteria for determining whether or not they have been achieved), this kind of ambitious cognitive goals are automatically likely to be judged as inappropriate on Laudan’s model. That is, if we have such goals, then Laudan does not provide a viable method for choosing between such methods.

(d) In inquiries we always use and appeal to many different standards: evaluative standards, methodological rules and logical inference rules. These standards may be related in various, more or less systematic, ways. To me it seems reasonable to think that, at least sometimes, whether or not a particular standard contributes to efficient goal-pursuit, partly depends on what other standards and rules are adopted at the time. For example, if one accepts some evaluative standard to the effect that it is rational to prefer the theory that best stands up to criticism, the result of applying this standard depends on what other standards and rules are used together with it. Such a standard would be pretty useless, for example, unless one observed some methodological rule to the effect that one should actively
try to criticize theories. Similarly, what results it leads to, depends on what standards of criticism are used — if these are so strong that all known theories are ruled out as unacceptable the standard in question becomes inapplicable; and if they are too weak and permit too much, it might lead to preferring theories that ought to be judged unacceptable.

For reasons such as these, it may at least sometimes be difficult to isolate particular standards and their contributions to goal-pursuit, in the way Laudan’s empirical tests of ends-means connections require.

(e) Perhaps most importantly, the application of Laudan’s model of empirical testing of methods (including evaluative standards), requires that methods and goals are clearly distinguished from each other. Otherwise there is no point to it. As I indicated in the previous section, evaluative standards and goals are often not distinct in this way. This is particularly so in cases where standards function as specifications of goals. To the extent that this is so, Laudan’s meta-methodology at least needs to be supplemented by an account of how different ”complexes” of specifying-standards-and-specified-goals can be rationally evaluated.

Laudan considers another objection to his account. The objection draws on a demand for justification (Laudan is himself a justificationist), and argues that Laudan’s proposal leads to an infinite regress: according to Laudan, a methodological rule is evaluated and justified by testing it empirically. That test presupposes some rule of testing, which must itself be justified by drawing on yet another rule, and so on.

Laudan’s answer to the objection is that ”we can avoid the regress provided that we can find some warranting or evidencing principle which all the disputing theories of methodology share in common. If such a principle — accepted by all of the contending parties — exists, then it can be invoked as a neutral and impartial vehicle for choosing between rival methodologies.” (Laudan 1996, 135) And, according to Laudan, there is indeed such a rule, a rule that he claims ”has promise as a quasi-Archimedean standpoint”. It is an inductive rule which he formulates thus:

(R1) If actions of a particular sort, m, have consistently promoted certain cognitive ends, e, in the past, and rival actions, n, have failed to do so, then assume that future actions following the rule ’if your aim is e, you ought to do m’ are
more likely to promote those ends than actions based on the rule 'if your aim is e, you ought to do n.' (Laudan 1996, 135)

I shall make a few brief comments on this: (1) The rule is rather weak: it applies only to clear-cut cases in which one method "consistently" succeeds while another one does not. (2) Finding a purportedly undisputed rule like this one does not constitute an answer to the objection. It does not show that the rule is justified in a way which stops the regress of justification. Instead Laudan changes the subject, and talks about agreement and the resolution of disputes. (3) Not all philosophers in the field of scientific methodology would accept (R1). Popper and other critical rationalists would, contrary to what Laudan argues (Laudan 1996, 136), not accept such an inductive rule.

Concerning my third remark, it should be added that a critical rationalist might indeed accept a very similar standard of testing, for example one which is formulated in terms of corroboration and which does not permit an inference from past success to the likelihood of future success. Such a rule might in some (or perhaps even most) cases give the nod to the same methodological rules as Laudan's inductive rule would. Whenever that is so, a dispute about the adequacy of competing standards or rules might be settled in a non-question-begging way without the existence of any shared "quasi-Archimedean" standard for settling disputes. It is the demand for justification which leads to a regress and to the search for a foundational standard. The justification problem is not solved by finding an undisputed rule, or rational consensus. And a rational consensus about some standard does not require any shared undisputed rule, much less a foundational one.

So concerning the rational resolution of disputes about standards, Laudan's account can be generalized in a way which dispenses with any "quasi-Archimedean" foundational standards — all that is required in such a case is that the different evaluative standards of the opposing parties lead to the same verdict (a case of domination, in which one methodological rule scores higher than the alternatives on both sets of standards).

---

10 Sankey has suggested this (Sankey 1997, 184 n. 20).
I think that Laudan’s account can (and should) be generalized in a similar and related aspect, too. I am thinking about the circumstances under which it is rational to accept a standard. Look at the properties Laudan ascribes to the rule R1: he writes that it is ”arguably assumed universally among philosophers of science” (which it is, arguably, not), that ”independently of the sociology of philosophical consensus, it appears to be a sound rule of learning from experience”, and that ”if (R1) is not sound, no general rule is” (Laudan 1996, 135). So Laudan expresses beliefs to the effect that R1 is adequate, that it is used and that there is no alternative to using it, and that it is not itself involved in any dispute over the adequacy of standards and rules. I think this can be generalized in a way which makes the reference to a ”quasi-Archimedean” standpoint superfluous. My suggestion goes along the following lines: when evaluating alternative standards, do so by using standards that are believed to be adequate, which have not been effectively criticized or shown to be problematic (so far), and that are not themselves at stake in the choice between the alternatives in question (question-begging).

As I said at the beginning of this section, I view these claims about the empirical testing of methods, as a contribution to meta-methodology — rather than to a theory about bootstrapping. This meta-methodology (”normative naturalism”) is a specific development of an element belonging to the more general reticulated model: the idea that standards and methodological rules can be evaluated as alternative means to cognitive goals. Here, in Laudan 1996, he proposes a particular analysis of methods (they are hypothetical imperatives asserting links between kinds of actions and kinds of outcomes), a particular empirical method for testing methods (they are to be tested against data from the past record of science), and a particular standard for making comparative evaluations of methods ((R1), or perhaps some stronger version of it).

The method and standard he proposes are rather weak, though, and allow for extensive revisions of (other) standards and rules. That is, they allow — within the constraints contained in this ”core” — for some measure of bootstrapping. The following passage expresses, in a general way, how Laudan himself views the bootstrapping character of his ”naturalistic” meta-methodology:
The naturalist uses the simple method of induction to 'bootstrap' his way to more subtle and more demanding rules of evaluation, which, in their turn, become the license for subsequent and yet more highly refined rules and standards. The virtue of this way of proceeding, and why it makes the foundationalist's search for deeper underpinning gratuitous, is that it is capable at any point of revealing its own flaws, if any. (Laudan 1996, 167)

I think this quotation reveals something important about the nature of Laudan's account of the evaluation and change of standards. In the first sentence, he summarizes his view of how standards are evaluated and improved. The "simple method of induction" (which probably refers to something like (R1)) is used to select other standards that have been successful in the past. Then one obtains a "richer" set of standards, which can be used at the next stage and whose application may lead to the adoption of still further standards, and so on. In this way, the set of standards may grow and become "more subtle", "more demanding" and "more highly refined". The picture seems to be that of a core, consisting of some inductive rule such as (R1), which is continually refined and added to, but which is itself kept constant through a sequence of changes (or perhaps a frame, within which changes are possible).

The second sentence, on the other hand, points to the possibility of criticizing and rejecting any previously adopted standards. Further applications of the procedure he proposes may lead to the detection of inadequacies in standards adopted at earlier stages of inquiry. Since Laudan claims that all standards are fallible and revisable, he must hold that even such standards as (R1), and "the simple method of induction", may themselves come to be criticized, shown to have "flaws", and be replaced. The core (or frame), that is, is revisable. It may be criticized and possibly rejected. But if the core with its specific standards were to be rejected, Laudan would not have very much of a general account left.

I want to emphasize that this is not a criticism of Laudan's meta-methodology. The point is rather that the higher degree of specificity of his meta-methodological proposals — which is what gives his naturalism what normative bite it may have — has the consequence that it cannot function as a general account of bootstrapping.
5.5 Briskman's problem-oriented bootstrap theory

The most fully worked out account of bootstrap rationality I have encountered, has been presented by Larry Briskman in "Historicist Relativism and Bootstrap Rationality" (1977). Like Popper, Bartley and Agassi, Briskman rejects justificationism in favour of a criticist view of rationality. With Popper, he proposes that instead of trying to prove or justify theories, we should try to criticize and improve them. When Bruce Hauptli discusses Briskman's theory, he somehow manages to miss this crucial point. Hauptli treats Briskman as a justificationist, which makes his treatment of Briskman's theory completely confused. Not surprisingly, he concludes that Briskman fails to justify the rationalist commitment to rationality (Hauptli 1995, 7-9).

This misunderstanding is especially unfortunate, since this is the only explicit discussion of Briskman's paper I have seen published.

In his theory about how to evaluate and revise standards, Briskman proposes that standards should be regarded as means to the goals of inquiries. This is something his theory shares in common with Laudan and others. However, Briskman attempts to incorporate this feature into a unitary account which is formulated in terms of the concept of problems.

With his bootstrap theory he is trying to develop a theory "which can acknowledge the historical 'relativity' (or variability) of all our knowledge (including all of our rational standards and intellectual aims) but which, nevertheless, avoids an historicist 'relativism'" (Briskman 1977, 520-1). He states that the problem, if one wants to avoid both ahistorical absolutism and historicist relativism is to develop a theory which can account for the possibility of improvement in both our rational standards and our intellectual aims. In effect, we want an epistemology which allows that we

---

11 Briskman refers to his own account as a theory of bootstrap rationality. For ease of exposition I shall adopt his term "bootstrap theory" while discussing Briskman's work, although it is not a theory in the sense I specified earlier in this chapter.

12 Briskman offers a (somewhat brief) criticism of justificationism on pp. 511-14.

13 Instead of Briskman's term "aim" I shall mostly use "goal". Briskman also introduces a distinction between standards for the evaluation of theories (etc.) and methodological rules for the improvement of theories (etc.). I use this distinction myself in this work, as I said in chapter 1.
can *learn to be more rational* — that we can learn to pursue better (or more rational) aims and can equally learn to pursue these better aims better (or more rationally). Any theory which can account for how this might be possible I shall christen, following Agassi, an epistemological bootstrap theory. (Briskman 1977, 521)

Briefly summarized, Briskman describes the bootstrap process as follows. Just like theories can be viewed as competing attempts to solve problems, standards and cognitive goals can be seen as *competitors* in relation to different kinds of *problem situations*. They can be competitors *only* in relation to such problem situations. Only when standards and goals can be viewed as competitors is it relevant to *compare* and *choose* between them. But if two standards or goals compete, the *problem situation* relative to which they are competitors *functions as a standard* against which they can be compared. That is, whenever there is need for comparing and choosing between rival standards or goals, there exists, in the form of a problem situation, a standard for evaluating them. Moreover, the problems against which standards and goals compete, arise in the very course of our inquiries and thus develop historically — that is, the standards we need for evaluating standards and goals are not themselves fixed, but develop as our inquiries progress from problem to problem.

I shall now proceed to spell this out in more detail.

**Problems and competitors**

Briskman begins by reminding us of some of the consequences of the criticist view of rationality — a view which focuses on critical evaluation and improvement of positions instead of on justification.

First, the (mere) possibility of doubt concerning something (a theory, standard, or goal) does not constitute an epistemological problem. The reason is that "doubt per se does not help us improve, since it doesn’t enable us to pinpoint any specific weakness in our current views which may then be improved upon — only concrete criticism can do this." (Briskman 1977, 521-2)

---

14 Briskman’s reference is to Agassi 1974.
Second, the point of rational evaluation is to determine preferences between actual, or existing, alternatives — by judging, e.g., that one theory or standard constitutes an improvement in comparison with another theory or standard: "the possibility of improved alternatives is not itself an improved alternative" — thus the possibility of, e.g., better or more adequate standards does not as such make the standards proposed so far problematical.15

Furthermore, an actual diversity of theories, standards or goals, does not always generate an epistemological problem, or, a problem of preference. Such a problem arises only when "the diverse views are alternatives or competitors." (Briskman 1977, 522) Only if theories or standards are alternatives which compete with each other, is it relevant to compare and choose between them.

What is it for theories to compete with each other, then? Briskman distinguishes between theories conflicting and competing with each other. For two theories to conflict with each other, it is sufficient that they are logically inconsistent. However, he claims, logical inconsistency is a necessary but not a sufficient condition for the theories to be competitors:

To see this, consider the fact that Darwinian evolutionary theory and nineteenth-century thermodynamics were shown to be inconsistent by Kelvin — yet there is no obvious sense in which they are competitors in the way that, for example, Einstein’s theory and Newton’s theory are competitors. (Briskman 1977, 522)

The point at which the two theories conflicted was concerning the age of the earth.

At this point Briskman introduces the concept of problems, and uses it to explain what it is for two theories to compete with each other: "two theories may be said to be competitors insofar as they can actually replace each other in our world view by offering conflicting solutions to the same, or highly overlapping, sets of explanatory problems" (Briskman

15 Cp. Laudan, who also insists that rational evaluation and choice, also concerning standards, is a comparative matter. According to both Briskman and Laudan, the rationality of choosing standards is a matter of choosing the best available ones, the best ones proposed so far. Of course, as Agassi reminds us, one would also want to require that the chosen standards should be (by some standard) good enough (Agassi 1974).
Einstein’s theory "offers solutions to a similar range of explanatory problems as does Newton’s theory" (Briskman 1977, 522) and can therefore replace it, whereas thermodynamics and evolutionary theory address largely different sets of problems. Accordingly, he suggests that theories "can only compete relative to an explanatory problem-situation." (Briskman 1977, 523)

Briskman applies this to standards and goals, too. He develops his bootstrap theory of rationality in two steps, dealing first with evaluative standards and methodological rules, and then with goals.\(^{16}\)

### Comparing competing standards and rules

Beginning with standards, he proposes that "standards and rules compete relative to an 'evaluational' problem-situation" (Briskman 1977, 523), and mentions two sorts of evaluational problems. He refers to them as "difficulties in the pursuit of an aim" and "problems of preference". Briskman treats these as two distinct kinds of evaluational problems, but I think it is more correct to treat problems of preference as belonging to the more general category of problems of goal-pursuit (problems of preference are problems of goal-pursuit).

In inquiries we pursue goals of various kinds, but we may of course fail to make progress toward them and run into difficulties of different sorts — we encounter problems of goal-pursuit. Briskman makes the following proposal: "Methodological rules (and their associated evaluative standards) can then be competitors insofar as they are suggested as alternative means for helping us overcome the set of difficulties..." (Briskman 1977, 536 n. 14)

---

\(^{16}\) At the end of the paper (Briskman 1977, 533) he says that a problem-oriented bootstrap theory which explains the rational evaluation and change of substantive scientific theories could also be given. In such a theory, scientific (explanatory) problems would function as standards of evaluation for comparing theories. Here Briskman refers to and acknowledges a debt to Hattiangadi’s work on problems, which was unpublished at the time (see Hattiangadi 1978 and 1979). This is an interesting approach, but it reveals a tension in Briskman’s position: he discusses how to improve general standards, while the problem-oriented view suggests that the relevant standards are instead particular problem-situations. This is indeed how Hattiangadi views the matter. See Hattiangadi 1983, which has the title "A Methodology Without Methodological Rules". Briskman tentatively proposes that we can evaluate theories either in terms of (general) standards or in terms of (particular) problems, and that when we use "standards to evaluate theories we are really using these standards as a surrogate for the problems they solve.” (Briskman 1977, 536 n. 14) The question is whether such a "reduction" of general standards to particular problem-situations will work.
encountered in aim-pursuit.” (Briskman 1977, 523) This is, of course, intimately related to the view of standards as means to goals.

Problems of preference, Briskman suggests, occur when existing standards are unable to determine a preference between existing competing theories:

as a result of the existence of articulated competitors at the level of theories we may be faced with a problem of preference which we cannot resolve with the help of our current standards. As a result, competing standards might be proposed as competing attempts to resolve this problem (Briskman 1977, 523).

However, one must add that such problems can always be "solved" by introducing some arbitrary, possibly crazy, standard (one might arrange a sort of lottery, for example, and decide that the theory whose name is written on the winning ticket is to be preferred). In an adequate account of bootstrapping, something more needs to be said about this. One important point is of course that such problems do not occur in isolation, and that a solution must be adequate to the problem-situation as a whole — most importantly, in relation to other existing standards and to the goals of the inquiry.

Furthermore, this kind of evaluational problem is also connected with the idea that standards are means to cognitive goals. It is, I think, a specific version of problems of goal-pursuit. To improve theories is a way of approaching goals, and evaluative standards have the function of determining preferences between theories (as better and worse) and select those theories which constitute improvements. So if a set of standards is, for example, so weak that it does not discriminate between theories and does not help us in judging improvement or the lack of it, then these standards fail to perform that function — they do not aid us in our goal-pursuit. Therefore this kind of problem is a species of the genus problems of goal-pursuit.

So, according to Briskman, standards compete and are evaluated relative to particular problems. The (evaluational) problem-situation functions as the standard against which evaluative standards are evaluated and compared, and an improvement of an evaluative standard (or a set of
such standards) consists in the fact that it enables "us to solve 'evaluational' problems which we were unable to solve on the basis of our previous standards and rules." (Briskman 1977, 524) This means that the standard for evaluating standards is something which itself develops in the course of inquiry. It depends on what difficulties we encounter in inquiry. I want to add that this, in turn, depends on what goals we are pursuing, what theories have been proposed, and what the current set of evaluative standards consists of.\(^\text{17}\) If this is so, then we cannot determine once and for all what this standard is (we cannot fix it), and we cannot know it in advance:

we cannot possibly know, in advance of both aim-pursuit and the articulation of competing theories, the concrete 'evaluational' problems with which our standards and rules will have to cope. But this entails that we equally cannot possibly know, in advance of aim-pursuit and specific competing theories, not only the ways in which our standards and rules might, in fact, compete, but also the ways in which they might, in fact, be in need of improvement, and hence the ways in which improvements in them might be possible. (Briskman 1977, 524)

So the problem-situation provides, according to Briskman, a non-arbitrary standard against which proposed evaluative standards can be evaluated and compared, and at the same time, this standard is not fixed but develops historically. As Briskman expresses it: "through pursuing aims and articulating competing theories we can actually learn to pursue aims and compare competing theories more rationally! Or to put it more colourfully — we can pull ourselves up by our own bootstraps." (Briskman 1977, 526)

**Comparing competing goals**

Briskman notes that one might object that "unless we can explain how our intellectual aims can improve — and how it is possible to nonarbitrarily\(^\text{17}\) As I said above, it is a weakness in Briskman's account is that he does not discuss the way in which previously existing standards may be involved in the generation of evaluational problems.
choose or evaluate them — we have not really overcome relativism without any appeal to an absolutism.” (Briskman 1977, 527) To meet this objection, Briskman proposes an extension of the bootstrap theory by adding to it an account of how goals are evaluated and improved.

As he did with standards, Briskman starts his discussion of goals by saying that the existence of a diversity of goals is problematic only if these goals are somehow in competition with each other.\(^{18}\) How can goals compete? Briskman’s answer is that sometimes goals function as means — as means for achieving other goals.\(^ {19}\)

He distinguishes between two different ways in which goals can be related as means to other goals: as ”sub-aims” and as ”surrogate-aims”.

Sub-aims of a more basic aim, he says, can be said to specify ”pragmatically necessary conditions which must be fulfilled in my pursuit of that aim.” His example of a sub-aim is the following: if ”my aim is to drive from Edinburgh to London; then a sub-aim of this aim might be to fill my tank up with petrol before I set out.” (Briskman 1977, 528) An example from the sphere of inquiry, might be that of an inductivistically inclined researcher who takes the collection of a body of data as a sub-goal of the more ”basic” goal of establishing an inductive generalization.

Surrogate-aims, on the other hand, ”may be loosely conceptualized as specifying pragmatically sufficient conditions which, if fulfilled, would amount to the fulfillment of some other aim.” To exemplify this, he writes: ”say that my aim is simply to get from Edinburgh to London. Then clearly the aim of driving from Edinburgh to London is not a sub-aim of this aim; rather it is a surrogate-aim of this aim. For driving from Edinburgh to London is getting from Edinburgh to London, as would by flying, taking the train”, etc. Thus ”a surrogate-aim is such that in pursuing it we are simultaneously pursuing another, usually more general, aim” (Briskman 1977, 528-29).

He gives another example of such a general goal: ”to explain the changes in the physical world around us” (Briskman 1977, 529). As examples of competing surrogate-goals of this goal he mentions ”(a) the

---

\(^{18}\) Briskman allows for pluralism with respect to non-competing goals.

\(^{19}\) This means that, as is the case with Laudan, Briskman offers an account of goal evaluation which is still situated within what might be called an instrumental view of rationality in inquiry. Thus, neither Briskman nor Laudan proposes any way of evaluating goals which is independent of instrumental considerations. They do not have any pure theory of value rationality.
aim of trying to explain the physical changes in terms of the desires, intentions, and aims of Spirits inhabiting the world; (b) the aim of trying to explain them in terms of things trying to actualize their potentia; (c) the aim of trying to explain them in terms of 'blind' deterministic laws.” (Briskman 1977, 538 n. 21).

Next, Briskman suggests that "intellectual aims can be in competition only if they can be viewed as surrogate-aims of some more 'basic' (or general) aim.” (Briskman 1977, 529)20 It might be the case, for example, that we start out on an inquiry with some goal (aim) which is so very general that initially "we can have little idea as to how to best pursue it.” This may make things difficult for us, and in order to get on with our inquiry we may "attempt to precisify the task we have set ourselves so as to actually gain more guidance in its pursuit”:

In this way we may be lead to articulate slightly more restricted aims which could function as surrogates for our initial aim — and such surrogate-aims can then [...] be seen as competing attempts to solve the problem of how we can better pursue our more general aim in the light of the continued difficulties which we have already encountered in its pursuit. (Briskman 1977, 529)

Myself, I think that "surrogate” is an unfortunate choice of term for what Briskman has in mind. His point is that if one achieves what he calls a "surrogate-aim”, then one also achieves the more general goal of which it is a specification. Briskman says that one pursues the more specific aim instead of pursuing the more general one, and this is probably why he chooses the term "surrogate": a surrogate is indeed something one uses instead of something else — instead of the real thing. However, when one buys a surrogate for something, one does not get the real thing. To have a cup of surrogate coffee is not to have a cup of coffee. Therefore, I think

20 To write "only" seems to me to be mistaken, and, actually, Briskman himself takes this back in footnote 22 (p. 538). There, he allows that sub-aims can indeed "compete qua means and so be evaluated in terms of their efficacy or efficiency in helping us achieve that aim of which they are sub-aims.” I see no reason not to think that sub-aims can be in competition. I think, however, that if sub-aims can compete in relation to a further goal, then they cannot be "necessary conditions which must be fulfilled in my pursuit of that aim”, although this is how Briskman proposes to explicate them (Briskman 1977, 528).
his terminology is a bit misleading. Myself, I shall from hereon instead use the term "specification-goal" (and substitute "specification-" for all occurrences of "surrogate-" in quotations from Briskman).

According to Briskman, then, the evaluation of proposed goals of inquiry works in much the same way as the evaluation of standards. His theory allows for pluralism with respect to goals; only if different goals compete with each other is it necessary to compare and choose between them. Just like standards, goals compete with each other relative to specific problem-situations, and are evaluated against these problem-situations. The kind of problem Briskman envisages, occurs when an existing goal is such that we do not know how to go about trying to achieve it, because the goal is too general or imprecise to guide or direct our inquiry. So he suggests that different specification-goals of such a more general (or basic) goal

compete as competing attempts to solve the problem of how we can better pursue our 'basic' aim in the light of continued difficulties (or problems) which we have already encountered in its pursuit. But this means that we immediately have a ready-made nonarbitrary standard for comparatively evaluating the competing [specification]-aims: for we can ask whether pursuing them either promises to, or actually enables us to, overcome or avoid these very difficulties. (Briskman 1977, 530)

Thus, Briskman claims, there are standards for evaluating proposed goals of inquiry, standards which can be used to criticize goals, compare competing goals, and make judgments about improvement of goals. These standards for goal evaluation consist of problem-situations, which depend upon developing stages of inquiries and the goals, standards, substantive theories and difficulties existing at particular stages. Such a standard "is not itself fixed (at some higher order, meta-aim level) but can rather develop under the impact of continuing aim-pursuit" (Briskman 1977, 530).

Briskman himself comments on the bootstrap character of his account of goal evaluation:
since aim-pursuit difficulties can function as a nonarbitrary standard for comparing competing (i.e., [specification-]) aims, historicist relativism is avoided; while since such aim-pursuit difficulties are not given all at once but rather emerge historically, this nonarbitrary standard is not somehow ahistorically fixed and known a priori but can rather be learned via aim-pursuit itself. Once again we can pull ourselves up by our bootstraps. (Briskman 1977, 531)

Further comments on Briskman

Briskman’s account does not satisfy the two desiderata I proposed a theory should satisfy in order to count as a real theory about the bootstrap process (since I claimed that no theory can satisfy them, this should come as no surprise). Briskman says something quite general about the bootstrap process and the role that the discovery of different sorts of problems play in it. But he does not specify in detail what kinds of considerations can motivate rational changes of standards and goals, and he does not give us any details about what sort of changes would be ruled out as irrational. That is, Briskman’s account satisfies neither the completeness nor the restrictions desideratum.

But that is as it should be: bootstrapping is an open-ended process. The idea behind the bootstrap view is precisely that we do not know today what inadequacies or difficulties we may be able to discover among our standards and goals tomorrow and how we may be able to improve them. By focussing on problems and the role they play in rational evaluation and change, Briskman is able to bring this out clearly.

How should one judge Briskman’s contribution to our understanding of what bootstrapping can amount to? The most important and most interesting suggestion he makes is the idea that the specific problems we encounter in inquiries — evaluational problems and difficulties of goal-

---

21 Briskman considers an objection that can be directed against his theory, an objection to the effect that it still contains an absolutistic element, viz. "the 'formal' knowledge that 'evaluational' problem-situations can themselves be used as standards of comparison." In reply, he writes that "we can discover this fact as an unintended consequence of the emergence of such problem-situations coupled with the attempt to solve the problems which have thus emerged. Even at this primitive stage we can pull ourselves up by our bootstraps!" (Briskman 1977, 537-38 n. 20)

Briskman might also have added that this piece of "formal" knowledge about goal evaluation is itself open to criticism and perhaps revision.
pursuit — can *themselves* be used as standards of evaluation. If that is so, then it means that in such cases we will be able to evaluate standards or goals even if we are unable to articulate any *general* standards that apply to the cases in question.

However, I think it is a weakness in Briskman’s account that he does not discuss the ways in which *general* (previously existing) standards of evaluation contribute to the generation of different kinds of problems, and how they put restrictions on what can be accepted as adequate solutions to new problems. Assume, for example, that we are faced with a problem of preference between two theories. We might introduce any one of a number of possible standards that would trivially "solve" this problem by leading us to prefer one theory over the other. But of course, most of these standards would themselves be unacceptable in the light of other standards that we already have (think of the absurd idea, mentioned before, of arriving at a preference as a result of arranging a lottery). For instance, do acceptable evaluative standards have to pick out a unique best theory to be acceptable, or is it enough if the standard in question can be used to discriminate between the alternatives that are actually in competition? And is it a virtue or a vice if the adoption of a particular standard tends to lead us to prefer more risky theories over less risky ones?

One would surely like to know more about how we can use our developing set of evaluative standards to *generate* those problems that, according to Briskman, provide us with opportunities of learning more about how to inquire rationally.

Before ending this chapter, I shall make some brief comments on the topic of goals.

### 5.6 Some more on goals

Briskman’s discussion of goal evaluation only concerns specification-goals, and how these can compete and be compared. There are, however, other ways in which goals can be related to each other, and in which goals may be revised. I shall discuss some of these, and comment further on "specification-goals".

(1) Specification-goals (what Briskman calls "surrogate-aims") are specifications of more general goals. To achieve a specification-goal is to
simultaneously achieve the more general goal it is a specification of. For example: to achieve a theory which gives deductive-nomological explanations of certain phenomena is to achieve the more general goal of explaining these phenomena; to achieve a theory which explains certain phenomena solely in terms of matter in motion and contact action is to achieve the more general goal of explaining these phenomena.

Competing specification-goals are, according to Briskman, competing attempts to solve the problem of how to achieve or approach a more general goal more effectively. A specification-goal can be evaluated in relation to the more general goal: proposing a more specific version of the goal may be judged a rational revision of the more general goal, if the specification-goal directs or guides inquiry better. It can also be compared with other specification-goals which are competing specifications of the same general goal — it might be judged to be a better guide than these, in that research along its lines is more successful, etc. It should be noted that a more precise specification of a general goal is here regarded as a revision of that goal.

Another thing that should be noted, is that what is often referred to as standards can be part of the specification of such goals. This was something I emphasized in the discussion of Laudan’s meta-methodology. Think of standards of explanation, for example. One might try to make a general goal of explaining certain things more specific by insisting that the explanations should be of some particular kind, i.e., by introducing standards of explanation into the goal specification (statistical or deterministic explanations, qualitative or quantitative explanations, deductive-nomological explanations). The same thing happens when goals are framed in epistemic terms, such as justified true beliefs rather than just true beliefs.

Another way of making the goals more specific is to propose specifications which might be said to be metaphysical. Examples might be to explain something in terms of contact action only, in terms of particles in motion or of fields of some sort, in terms of individuals or collectives, etc.

Thus, the critical evaluation and improvement of (specifications of) goals of inquiry may involve debate about the appropriateness of various

22 In the liberal sense of ”revision” that I use here.
standards of explanation or interpretation, etc., as well as about the appropriateness and heuristic value of metaphysical ideas.

There are, however, other ways in which goals can be related, and I would like to mention some of them here.

(2) One way in which one goal can be related to another, is as a part of a whole. Assume for example that one has the (general) goal of explaining the behaviour and properties of matter. Assume further that one knows that matter has various magnetical, electrical and dynamical properties. A possible goal is then to develop a theory which explains all these different characteristics of matter. It might turn out, however, that it is difficult to find a theory which explains all this (or even a promising line of research towards this goal). In response, one might propose more limited, or partial, goals: the goal of explaining magnetism, or of explaining electricity, for example. This would be an instance of what Shapere calls the "domains approach" (Shapere 1984).

Such partial goals are related to the goals of which they are parts, in a way which differs from how Briskman’s specification-goals are related to their more general goals. If one manages to develop a theory of electricity, one does not thereby have explanations of magnetism, falling bodies or the behaviour of gases. Neither do these partial goals compete with each other in Briskman’s sense. However, to adopt such limited goals may constitute rational revisions of more comprehensive goals, when the latter turn out to be difficult to pursue. The pursuit of a partial and more limited goal may prove a more manageable task.

(3) Another type of relation between goals can be exemplified by the relation between truth and increasing degrees of approximation to the truth (verisimilitude), or between theories known with certainty to be true and theories with a high degree of probability. Here, the difference can perhaps be expressed as one between more and less ambitious goals. One decides to aim, so to speak, for the "next best thing", if the initial goal is judged to be too difficult to achieve. If certainty or truth as such turn out to be unattainable, one can propose to aim instead for something more modest that one believes is achievable — such as probability or getting nearer to the truth. The relation is not one of part to whole. It also differs from the relation between specification- and general goals — certainty entails truth while probability does not, and a series of theories with increasing degrees of verisimilitude may all be false. The less
ambitious goals do not "include" the more ambitious ones. (Actually, these kinds of goals, where you don't automatically get the real thing, might perhaps be called "surrogates" for the more ambitious goals they replace.) Still, these less ambitious goals seem to be intimately connected with the goals they replace — after the goal revision, one would still be trying to do very much the same thing. I do not know how this continuity should be explicated, though.

Revisions of goals along the lines of (2) and (3) amount, one could say, to retreating from a more ambitious but seemingly unattainable goal, and to the adoption of goals which seem to be easier to pursue. However, it is also possible to raise the stakes and adopt more demanding goals (by seeking true explanations rather than empirically adequate theories or even just successful predictions, by demanding unification or reduction, or higher degrees of precision or probability, for example). It would be interesting to know more about the rationality of this kind of goal revisions, but I shall not pursue this further. These distinctions between different kinds of relations between and revisions of goals is just intended to show that there is more to cognitive goals than Briskman says in his paper, and to point out that much work remains to be done on the topic of the rationality of cognitive goals.

5.7 Conclusion
Laudan and Briskman offer accounts of bootstrapping that are both partial and hypothetical. They are partial in that they specify (in a general manner) some ways in which standards can be rationally evaluated and revised, but they do not give exhaustive accounts of what can be rational revisions. It is part of the general bootstrap view that new standards, and thus new means for evaluating and revising standards, may come to be rationally added to the set of standards we have at any given point. As far as I can see, their accounts are consistent with the possibility of there being other ways of evaluating and revising standards rationally, besides the ones they present.

Their accounts are also hypothetical, in that they contain certain assumptions which are themselves open to challenge. For one thing, assumptions to the effect that certain standards are adequate are built into their accounts. Laudan, for example, takes rationality to require
justification. Accordingly, he holds that if it is to be rational for an agent to accept a standard, the agent must have some positive reason for thinking that the standard in question is a means for achieving his goals. Similarly, he holds that an agent is rational in adopting a goal only if he already has reasons for thinking that the goal is achievable.

Although Laudan himself does not say so explicitly in the case of these standards, it follows from his claims about the revisability of all standards, as well as from the very idea of bootstrapping, that these standards should also be held open to criticism and revision. Indeed, Briskman explicitly rejects the idea that rationality requires justification, and he accepts neither of the two standards above. According to Briskman it can be rational to adopt a goal or a standard without having any reasons that make one justified in thinking that the goal is achievable or that the standard is a good means to one's goals.

To say that their accounts are partial and hypothetical amounts, however, to the same thing as saying that they do not meet the completeness and restrictions desiderata and are thus not real theories about bootstrapping (in the sense introduced in section 5.1). But since it is a consequence of the bootstrap view itself that there can be no such theory, this was entirely to be expected. Both Laudan and Briskman still make contributions to a richer understanding of ways in which bootstrap operations can allow for rational evaluation and revision of standards.

Nevertheless, their accounts of bootstrapping have a regrettable limitation in common. In Laudan’s reticulated model, previously accepted standards will be involved in evaluating all the kinds of changes of standards and goals that Laudan describes. This is something he does not discuss explicitly. Similarly, when Briskman writes about how problem-situations can function as standards for evaluating standards and goals, he does not discuss how previously existing standards function in generating problems and in imposing restrictions on rationally acceptable solutions to such problems. In this chapter I have tried to identify limitations such as these, although I have not tried to supplement their accounts in any detailed way.

There are some further issues concerning rationality which remain to be dealt with.

One of them has to do with an assumption about rationality and standards that both Laudan and Briskman make. They both think of the
adequacy of standards as being a matter of how well they help us in achieving the goals of inquiries. This is the instrumental view of standards, a view they share with the majority of philosophers of science. This instrumental conception of adequacy is built into their accounts of bootstrap rationality, but it seems that this conception is itself something that can be discussed and challenged. What are the consequences for a normative understanding of rationality in inquiry if even the idea of adequacy — and thus the idea of the improvement of standards consists in — is open to criticism and revision?

A related question is also raised by the view that rationality has to do with goal-pursuit, and the instrumental conception of the adequacy of standards. Both Laudan and Briskman have presented ways in which proposed goals of inquiry can be rationally criticized, and it seems that there can at least be rational restrictions on what goals inquiring agents adopt. But what if different agents adopt different goals, and if different standards are adequate for the goals in question? Does that make the concept of adequacy, as well as the concept of an absolutely rational action, unacceptably relativistic?

Furthermore, according to the bootstrap view discussed in this chapter, what it is rational for an inquiring agent to do, is something that is at least in part relative to a certain stage of inquiry and discussion, and to the standards the agent accepts as adequate. In Agassi’s words: "What was rational enough yesterday need not be rational enough today" (Agassi 1974, 414). The idea of a rational action is intimately connected with the idea of conformance to adequate standards, whereas the bootstrap view says that it may be rational at one point to use a standard which is later discovered to be inadequate. The view of rationality which has emerged so far in this work, raises questions about whether, and how, this revisability of standards can be reconciled with an objectivistic conception of adequacy and with a non-relativistic theory of rationality.

In the next chapter I try to answer questions such as these and make it clear exactly how the normativity of rationality in inquiry and of the adequacy of standards, is to be reconciled with the revisability and relativity of rationality, as well as with the idea that rationality has to do with goal-pursuit.
6

RELATIVE AND ABSOLUTE RATIONALITY

6.1 Introduction
In this chapter I discuss issues concerning the normativity of rationality in inquiry — how the normativity of standards, and of evaluations in terms of rationality, is related to the view of rationality proposed in earlier chapters. The problem is that of reconciling the change and diversity of standards and rules with a normative and non-relativistic conception of rationality.

I have argued that we are fallible with regard to standards and methods of rational inquiry — that we cannot rule out the possibility that the standards and rules we employ are less than fully adequate or even mistaken. I also argued that there are no good reasons for treating any of our standards or rules as uncriticizable and unrevisable, and proposed instead that the most rational policy is to treat all standards and rules of rational inquiry as open to criticism and improvement. In chapter 5 I discussed and tried to characterize the bootstrap view of rationality, an abstract idea of how standards and rules can change rationally which does not require that there is any level of fixed unchanging standards to evaluate such changes.

From the presentation of the bootstrap view in the previous chapter, it is clear that what it is rational for an inquiring agent to do, is something that is at least in part relative to the standards the agent takes to be adequate, and to a certain stage of inquiry and discussion. The bootstrap view allows that at a particular stage an agent may rationally use a standard which he later comes to discover is inadequate and revises. On the other hand, the idea of a rational action is intimately connected with the idea of conformance to adequate standards. Precisely how is the connection between the adequacy of standards and the rationality of actions to be understood? And if what it is rational for an agent to do
depends on what standards he accepts, does not this lead to rationality relativism?

I argue that if rationality is to be capable of guiding the actions of inquiring agents, what an agent rationally ought to do in a certain situation must be relative to the standards that he accepts. An agent may be acting rationally even if it is the case that his action violates adequate standards.

The concept of the adequacy of standards, and the relation between the adequacy of standards and the rationality of actions, are further discussed. So are certain complications with the instrumental conception of adequacy.

I conclude the chapter by showing that the account of rationality in inquiry presented in this work does not entail rationality relativism.

6.2 Normativity and availability, change and diversity

The concept of rationality discussed in this work is a normative concept, and methodological rules and standards of rational evaluation have to be capable of guiding, in some sense, the actions of inquiring agents. The normativity of standards involves some sort of claims to the effect that the standards we appeal to are in some sense adequate, and that the decisions or preferences they lead to are ones that we rationally should or ought to make or have. If rationality, through the adoption of standards of rational inquiry, is to be appropriately action-guiding, rationality must somehow be accessible to us. Normative rationality, what we rationally ought to do or how we should act in conducting inquiries, must have a fairly intimate connection with what we believe to be rational, with what we at some point accept as adequate standards of rational evaluation and as adequate methodological rules. To be properly action-guiding, normative rationality must in a sense be relative to the standards that inquiring agents have adopted or believe to be adequate.

At the same time, standards can change over time, and sometimes different agents or groups of agents accept different standards and disagree about which are the adequate ones. The arguments of the earlier chapters have been to the effect that this is as it should be. We are fallible

---

1 See Foley 1988 and 1993. This is also closely connected with the distinction between internalism and externalism in epistemology. For an overview, and for arguments against externalism, see e.g. Pollock 1986 and Haack 1993. Also, see Nozick’s distinction between theories of “best action” and theories of “rational action” (Nozick 1993, 65).
with regard to standards of rationality and may be mistaken about which ones are adequate, and we may be lucky enough to find criticisms which allow us to discover such inadequacies and hopefully improve our standards. All evaluative standards and methodological rules, I have argued, should be held open to criticism and revision.

So on the one hand we cannot reduce rationality to the standards accepted and used by particular agents at a particular time. That would amount to rationality relativism. On the other hand, what is rational for a given agent in a particular situation must at least partly depend on that agent's standards. One cannot take account of and be guided by standards one has never heard of, and as one tries to conduct inquiries rationally one has no option but to use those standards one has and takes to be adequate.

If rationality is to be action-guiding, it must thus to some extent be relative to the standards adopted by different agents, but it cannot be identical with or reducible to them. So I will argue. This chapter deals with the problem of giving an account of normative rationality which allows it to be relative without leading to relativism. The central difficulty is how to reconcile the idea of non-relativistic normative rationality with the possibility of diversity and change concerning standards of rational inquiry.

6.3 Normativity and adequacy

The normative nature of rationality in inquiry is connected with the idea of adequacy — the idea that standards can be, or fail to be, adequate.

When we claim that for example a decision is rational or irrational, by arguing that it conforms to or violates some evaluative standard, we are not just describing it. We are not simply stating that the decision in question conforms or fails to conform to a given standard. Instead we are making a normative evaluation, or a recommendation. And when such an evaluation is made by appealing to particular standards, a claim is involved that the standards appealed to are in some sense adequate, or correct, and that they are such that we rationally ought to conform to them (or at least not violate them).
How is this adequacy or correctness to be understood, then? What is it for a standard to be adequate or inadequate, or for one standard to be more adequate than another one?

Myself, I think that the adequacy of evaluative standards and methodological rules is a matter of the extent to which they are suitable means for achieving the goals of inquiry. This is the majority view among philosophers. According to this view, standards are instruments, and their adequacy consists in being efficient means in cognitive goal-pursuit. This provides a rationale for adopting and using standards — we use them in inquiry because we believe that helps us achieve our cognitive goals. It also gives some content to the idea that one standard can be more adequate, or better, than another one: it is more adequate if it better enables us to achieve the goals of inquiry.

However, I want to add some reservations and qualifications.

First of all, this is not a conceptual truth about the adequacy of standards. I think that standards should indeed be seen as means that we use in order to pursue cognitive goals as efficiently as we can, but this instrumental conception is of course open to challenge. This idea of adequacy can itself be discussed, and I would not want to rule out the possibility that there might exist good reasons for different conceptions of adequacy. It is for example at least intelligible, I think, to claim that some standards simply are such that one rationally ought to respect them even though they have nothing to do with goal-pursuit. Perhaps something like a "deontological" conception of cognitive rationality could be formulated, as D'Agostino has proposed.²

Second, even if the adequacy of standards is ultimately a matter of how well they function in enabling us to pursue the cognitive goals of inquiries, to make this claim is not very informative. A particular standard may contribute to goal-pursuit in general by fulfilling a very specific function. For different groups of standards it is possible to specify what it is for them to be adequate. This can be done with varying degrees of precision. Think of logical inference rules, for example. For a logical inference rule to be adequate is precisely to be valid, that is, truth-

---

² See D'Agostino 1990 and 1995, and the discussion in Schulte 2000. D'Agostino writes about rules of scientific method that it is "rational deontologically, expressively, that scientists conform to these rules — not because of any anticipated benefits (though these may result), but, instead, because of the intrinsic qualities, relative to their commitments, of decisions in conformity with these rules." (D'Agostino 1990, 35)
preserving. A method of statistical sampling is instead adequate to the extent that it leads to nonbiased samples. The idea of a nonbiased sample may not be very exact — not as exact as the idea of deductive validity — but it is much more specific than the idea of just being an appropriate means to cognitive goals. In cases such as these it is often possible to specify in a more or less precise way what the adequacy consists in, and one can do this without saying anything about the goals of inquiry. Still, I think that the reason that these kinds of adequacy are relevant and normatively interesting in the first place, is precisely that deductive reasoning and statistical sampling serve important functions in a wider context — the context in which we inquiring agents try to find the truth, make predictions, and so on. That is, they are important for rationality because they are, after all, important for agents as they pursue their cognitive goals.

Third, as I suggested in sections 5.4 and 5.6, standards can be partly constitutive of goals. This is the case when goals are specified in terms of standards. The goal of true explanations, for example, may be specified in terms of some standard of adequate explanation, or in terms of some epistemic standard. In such cases it is clear that the standards in question cannot be regarded as externally related to independent cognitive goals.

### 6.4 Rationality as relative to agents in their situations

It is generally acknowledged that what is rational for an agent depends on the resources that are available to the agent in the situation in which he finds himself. The rationality of an agent’s action or judgment in a particular situation may depend upon the information he has, what his goals are, what alternatives are open to him, and so on. What rationality normatively requires of an agent is to some extent relative to certain aspects of his situation. Nicholas Rescher expresses it as follows: "Rationality is a matter of deliberately doing the best one can with the means at one’s disposal — of striving for the best results that one can expect to achieve within the range of one’s resources — specifically including one’s intellectual resources.” (Rescher 1988, 1-2)

In any instrumental view of rationality, the question of what it is rational for an agent to do depends on what goals he has. Similarly, it is generally held that rational belief or rational theory acceptance depends
on the information, or evidence, that is available to the agent at the time. I would like to express this by saying that rationality is both goal-relative and information-relative, and I take it that I am not saying anything controversial here.

The goal- and information-relativity of rationality constitute two ways in which the rationality of an inquiring agent depends on the situation he is in, and on his appreciation of that situation. I would also like to add that what it is rational for an agent engaged in inquiry to do — for example, whether he should accept or reject a theory, or whether he should prefer one theory to another one — also depends on his awareness of available arguments and on his awareness of alternatives.

These different ways in which rationality in inquiry may be said to be relative to the particular situation of an inquiring agent at a particular stage of inquiry, I shall bring together under the term "agent-relative rationality". They all have to do with the fact that, if rationality is to be able to guide inquiring agents, then rationality cannot require of agents that they do the impossible, or that they do more than can reasonably be expected of them given the situations they are in. Ought implies can. One cannot, for example, take account of information or evidence that is not available at the present time. An agent may rationally try to achieve a goal, which he only later discovers is unachievable. An example might be Hilbert's goal of proving the consistency of arithmetical theories by finitistic means, which Gödel proved was impossible. In deciding which theory is the best one in some particular field, one cannot consider theories one has never heard of or which have not even been formulated by anyone yet. And arguably, without being irrational an agent may fail to realize that two statements are inconsistent with each other or that one statement is entailed by another, if the argument that shows this to be so is sufficiently complicated. As an example, think of Frege, before Russell showed his set theory to be inconsistent.

To be rational consists in doing the best one can, or doing sufficiently well, with the resources one actually has. The point with this idea of agent-relative rationality, is that the situations in which inquiring agents find themselves are typically such that these resources are limited: they do not have all the relevant evidence, they are not able to think of or imagine all logically possible alternatives, they fail to realize complex logical relationships, and they may aspire to achieve things that eventually turn
out to be unachievable. What an agent rationally ought to do in such a situation must be something that is possible and feasible in spite of such limitations. Thus, given that a piece of refuting evidence has not been made public yet, an agent may be rational in accepting a theory which that evidence refutes; given that he has not yet discovered or seen a proof he may rationally refuse to accept a mathematical statement as a theorem; he may rationally favour a particular type of theory as the best explanation of a class of phenomena, given that he has not discovered or been presented with the new theory which explains them even better, and so on.

My main thesis in this chapter is that in an important sense rationality — what an agent rationally ought to do — is also relative to the standards accepted by inquiring agents. An important aspect of the situation of the inquiring agent consists of the standards of rational inquiry the agent accepts. This idea was introduced in chapter 1, and I shall try to develop it further here.

I am not thinking merely of the trivial fact that an action can be described as conforming to or violating the standards a particular agent accepts. Rather I make the normative claim that an agent rationally ought to conform to his own standards, and I shall also argue that this can be what he ought to do even if that means that the agent is thereby violating what are in fact adequate standards.

The standards possessed by an agent may be regarded as resources at the disposal of the agent: the standards he has enables him to evaluate theories and arguments, to devise strategies for testing and improving theories, and so on. To be able to inquire at all, an agent must have some standards or other, but some of the standards of an agent may, without his knowing so, be more or less inadequate. Such an agent may furthermore be in such a situation that he is not aware of the arguments and information that would reveal the inadequacy of his standards, or it may be that he does not know of any alternative standards that might be more adequate than his present ones — he may be, that is, in a situation in which from his own point of view he does not have any reasons for suspecting that his standards are inadequate. In such cases, I will argue, an agent who judges and acts on the basis of an inadequate standard may be acting rationally although his standard is inadequate and leads to decisions which by objectively correct standards would be ruled out — given that he does
accept the standard in question and is not in a position to realize its inadequacy.

I will thus be proposing that rationality be understood as relative not only to the information, goals, awareness of alternatives and arguments that are available to inquiring agents, but also to the agent’s own standards of rational inquiry. This I propose to refer to as the ”standard-relativity” of rationality, or as ”standard-relative rationality”. In the following section I will discuss this concept and contrast it with what I call ”absolute rationality”.

6.5 Standard-relative and absolute rationality
An inquiring agent appeals to or tries to act in conformance to his methodological rules, evaluative standards and logical inference rules as he conducts his inquiry. To be able to make any evaluations at all, he must — explicitly or implicitly — appeal to some standards or other. And they must be his own standards, ones that he himself accepts as adequate and as having a normative force. The agent, in trying to come to a rational decision, or to determine his preferences rationally, has no option but to use those standards he actually takes to be adequate and try to conform to these.

I am employing Pollock’s distinction between conformance to and accordance with standards (see section 1.6). For an action to be in accordance with a standard it is sufficient if the action would be permitted (or perhaps required) by that standard, whether or not it is a standard that the agent in question is himself appealing to. That an agent’s action conforms to a standard means that the agent is actually guided by that standard as he acts.³

³ This can be compared to C.D. Broad’s threefold distinction between ”action in accordance with, action on, and action for a principle” (Broad 1930, 119). Acting on a principle in Broad’s sense would be conforming to it in Pollock’s sense. So would acting for a principle be. But in Broad’s terminology, only actions ”done on a principle which is a categorical imperative” are actions for a principle. Pollock’s ”conformance to” applies to both hypothetical and categorical imperatives.

Perhaps it would also be fruitful to distinguish between two kinds of conformance to: conforming to a standard because it is taken to be adequate, and conforming to a standard while remaining neutral about whether it is adequate or not (for example, for the purpose of testing the consequences of using a standard, or in order to please one’s teacher).
What judgments an agent is able to make about what is rational, is limited by and depends on what standards he actually possesses (in a certain situation, at a particular stage of inquiry).

I shall call evaluation and other action in conformance to standards that are fully adequate "absolutely rational." "Absolute rationality" as I will use it stands for acting, deciding, etc. in conformance to standards that really are perfectly adequate. "Absolute" here simply means "not relative" — that is, not relative to what standards inquiring agents actually accept as adequate, not relative to what agents believe about the adequacy of standards. That an evaluation, for example, is a case of or an expression of absolute rationality, means that the evaluative standard in terms of which the evaluation is made really is adequate. Whether a standard is adequate or not, is not relative to what agents believe about its adequacy. It should be noted, though, that absolute rationality does depend on what standards an agent accepts, since it requires conformance to adequate standards, which in turn means that the agent in question is actually using the adequate standard.

By saying that an evaluation or other action is "standard-relatively rational" I mean, as a first approximation, that in making or performing it an agent is conforming to the standards he himself accepts as adequate. "Standard-relative rationality" thus stands for acting, deciding, etc. in conformance to one's own standards — whether these really are adequate or not. In this standard-relative sense, rationality is a matter of living up to, or not violating, one's own standards of rational inquiry.4

Standard-relative rationality is one aspect of the agent-relativity of rationality. What actions are rational depends on something about the agent performing the action: it depends on what set of standards the agent in question accepts. Thus if there is variation in what standards different agents accept, then what actions are standard-relatively rational for particular agents may vary, too. So one action, for example the decision

---

4 The distinction between standard-relative and absolute rationality can be compared with a distinction Jarvie and Agassi have made between three kinds of rationality. The first is "(1) weak rationality, pertaining to an action as explained by given goals". Then there are two stronger kinds, "pertaining only to goal-directed actions based on beliefs held to be rational by some standard." They are "(2a) relatively strong rationality where the standard of rationality evoked is that of one's own community, and (2b) very strong rationality, which pertains to the highest standards of rationality known anywhere" (Jarvie and Agassi 1987, 392). Also, compare Goldman's distinction between "strong justification" and "weak justification" (Goldman 1992, ch. 7).
to prefer theory A to theory B, may be standard-relatively rational for one agent but not for another.

Standard-relative rationality is to be understood as a genuinely normative concept, I will argue. A grasp of this normative concept is also important for understanding the rationality of changes of standards of rational inquiry.

However, standard-relative rationality is of course not all there is to normative rationality. Concerning any standard one can ask whether it really is adequate, and whether one rationally ought to evaluate by that standard.

An action such as evaluating one theory as better than another one, is standard-relatively rational for an agent to the extent that it conforms to the evaluative standards that agent accepts as adequate. But the fact that an agent accepts a standard S and believes that it is adequate, does not mean that it really is adequate. The agent may be mistaken about this.

Standard-relative rationality is rationality in a qualified form. It has to do with how I rationally ought to act, given that I accept certain standards as adequate.

Absolute rationality, in contrast, lacks this qualification. Instead it stands for how I rationally ought to act according to standards that are in fact fully adequate.

If one wishes to, one could say that standard-relative rationality is a matter of how I rationally ought to act given what I at present take to be adequate standards, and that absolute rationality is an idealization which stands for how I rationally ought to act if I were apprised of what standards are actually adequate.

There may thus be a discrepancy between absolute and standard-relative rationality. Of course, if in performing an action I am conforming to my own standards, and if these standards actually are adequate, then my action will be both standard-relatively and absolutely rational. They coincide when the agent’s standards are adequate. But we may be mistaken about what standards are adequate. I may conform to my own standards and thus be standard-relatively rational, but fail to be absolutely rational because the standards I appeal to are somehow flawed. It is this case, in which an agent meets the requirements of rational inquiry as he himself sees them but fails to be absolutely rational because some of his standards are objectively inadequate, that is of most interest to me.
There are two more permutations. One is when the agent’s action fails to be rational both standard-relatively and absolutely, that is, when his action violates both his own standards and the ones that are in fact adequate.

The other and final combination is when an agent’s action violates the standards he himself accepts, but happens to accord with what the standards that are in fact adequate would recommend. The action is then not standard-relatively rational. Is it absolutely rational? It is not. The action is the one that is required or permitted by adequate standards, so the agent is in a sense doing the right thing. The action happens to be in accordance with adequate standards, but the agent is not conforming to these. Furthermore, one aspect of his action is that he is violating his own standards, and this, I believe, is never rational. Indeed, I think the following standard is an adequate one: ”It is irrational to violate one’s own standards of rational inquiry”. What this means, is that standard-relative rationality is a necessary condition for absolute rationality. Absolute rationality requires not only that one’s action accord with adequate standards; it also requires that those standards be ones that one accepts as adequate and tries to conform to (that one’s action is guided by those standards).

This is part of the reason for regarding standard-relative rationality as a genuinely normative concept: it is always irrational for an agent to knowingly violate his own standards, and he always rationally ought to respect what he himself takes to be adequate standards.5

That evaluative standards and methodological rules are supposed to be action-guiding also provides a reason for regarding standard-relative rationality as normative. An agent, in trying to be rational, has no option but to try to act so as to meet the standards he himself takes to be adequate. As long as he takes this particular set of standards as adequate, this is the best he can do.

Let us look at an abstract example. Suppose that I am comparing two theories T1 and T2, trying to decide which of them is the best one.

---

5 One might, of course, without irrationality, temporarily put a standard on ”suspension” and consciously violate it in order to see what happens. Similarly one might, hypothetically so to speak, examine the consequences of a standard one does not accept, by trying to act on it and seeing what results. But these cases are atypical. Furthermore, in evaluating these consequences one must appeal to some standards one does accept and concerning which one does not suspend judgment.
Suppose there is a standard $S_1$ that is relevant for comparing the two theories, and which I think is adequate and such that the rational decision ought to be made by appealing to $S_1$. According to $S_1$, $T_1$ is better than $T_2$, so I prefer $T_1$ to $T_2$. Given that I accept $S_1$ this is then standard-relatively rational. Suppose further that $S_1$ is nevertheless in fact inadequate, and that instead a different standard, $S_2$, is adequate. According to $S_2$, which is fully adequate, $T_2$ should be preferred to $T_1$. But I have never heard of $S_2$, and I do not know that there is anything problematic about my $S_1$. What should I do in this situation? I cannot make my decision in the light of $S_2$, since I am not even aware that there is such a thing as $S_2$, or that $S_2$ is as a matter of fact adequate. Of course, it would be good if I knew that $S_2$ is adequate and made the correct decision in favour of $T_2$ — but in the example I do not know this. The best I can do in this situation is to make my decision conform to the standard I actually believe to be adequate. $S_1$ is in this case the only guide available to me, and to try to act in conformance to it is the best I can do. Standards such as $S_2$, which I do not know about, cannot function as guides for me in my cognitive activities — which is how standards of rational inquiry are supposed to function.

Since we are all (as I argued in chapter 2) fallible with regard to standards and rules, we cannot rule out the possibility that some of our standards or rules are inadequate. In certain instances we may fail to be absolutely rational, and unless we are able to detect and remedy such inadequacies by revising our standards, then in those instances we will at best be standard-relatively rational. Whether our standards really are fully adequate or not, the best we can do in the meantime is to try to meet those standards we do believe to be adequate.

I have already said that standard-relative rationality is a necessary condition for absolute rationality. The concepts are related also in the following way. Accepting a standard is accepting it as *adequate*, as something one normatively ought to conform to. Therefore accepting a standard means regarding it as an expression of absolute rationality. An agent who acts on the basis of his own standards implicitly claims that these standards are not merely *his* standards but that they *are* adequate. He implicitly claims to instantiate absolute rationality. He may of course fail to do so, if his standards are inadequate.
I have explicated standard-relative rationality as action that conforms to the standards of an agent. Given that an agent accepts standard S and his action conforms to S, he is standard-relatively rational. What I have emphasized so far is that it need not be automatically irrational to accept a standard which is inadequate, and that acting on the basis of such a mistaken standard may on the contrary be normatively rational. But might it not be irrational for the agent to accept the inadequate standard in the first place?

Yes, we can be rational and irrational with regard to our standards, too. Indeed, that has been a major theme in previous chapters. The concepts of standard-relatively rational and absolutely rational apply also to the acceptance of evaluative standards and methodological rules themselves. I may for example hold a certain meta-standard MS according to which I am rationally permitted to accept a certain standard S and thus be standard-relatively rational in accepting S. At the same time MS may itself be an inadequate meta-standard, and accepting S may fail to be absolutely rational. It is also possible that my accepting a certain standard S is standard-relatively irrational for me in the light of other standards that I also accept, so that I rationally ought not to accept S.

As an example, assume I have just been informed of some new empirical hypothesis, and am now considering whether to accept it or reject it (or take some other attitude to it). It happens to be a hypothesis which, if I accepted it as true, would somehow threaten a central part of my world picture. It could be a hypothesis about genetic differences between ethnic groups which does not fit in with my political outlook, or a hypothesis which contradicts a theory I am trying to build a career on. Anyway, I hope it is not true. Further assume that in considering the merits of this hypothesis I apply to it some very demanding standard, a standard requiring that there be conclusive proof for a theory to be acceptable. The hypothesis is of course nowhere near conclusively proven. I conclude that this worrying hypothesis does not even merit interest and further investigation, and so it does not worry me any more. However, otherwise I know that such a standard is excessively demanding and would lead to total skepticism, and it is not a standard of evaluation I normally use. By my usual standards, the hypothesis would instead look pretty well. Let us also assume that at least implicitly I hold a metastandard of impartiality, which rules out using "double standards" in
this way. In this example, my evaluation of the abominable hypothesis does conform to the standard I decided to apply to it. But I am clearly standard-relatively irrational in applying that standard and in dismissing the hypothesis: using that standard because it suits me on this occasion means that I violate my other standards.6

But once again, as was noted in the chapter on bootstrap processes, to be able to evaluate a standard \( S_1 \) as problematic, the agent needs to use some other standards — and these other standards must be ones that he himself accepts as adequate. The idea that standard-relative rationality is normative is actually a natural counterpart to the bootstrap view of rationality discussed in chapter 5 — and indeed of any view of rationality according to which it is possible rationally to change standards and rules.

### 6.6 Doing one’s subjective rationality duty and doing what is advisable

But is there not something deeply unsatisfactory about saying that an agent that acts in conformity with standards that are inadequate is doing what he rationally ought to? Evaluating his action from an external perspective, would one not want to say that, to the contrary, what he ought to do is what adequate standards require? I think the situation can be made clearer by keeping two kinds of normative questions apart.

Gibbard makes a distinction between rationality and advisability. The difference between what is rational for an agent and what is advisable, he formulates in terms of information the agent has or lacks. Rationality, he says, ”often consists not of using full information, but of making best use of limited information. Acting in full awareness of all relevant facts suggests not rationality, but something more like ’advisability’. Whereas rationality is a matter of making use of the information one has, advice can draw on information the advisee lacks.” (Gibbard 1990, 18). Gibbard writes about information, or awareness of facts, but this can also be applied to the standards agents use in inquiry. One could perhaps speak of agents having or lacking information about the adequacy of standards.

In relation to standards, I shall make a similar distinction between doing one’s rationality duty and doing what is advisable. To see what it amounts

---

6 The example is a case of (perhaps extreme) self-deceit. I do not mean to suggest that failure to be standard-relatively rational automatically means that one is deceiving oneself.
to, let us look at an example. Imagine an agent who is contemplating whether or not to accept a new theory which has been proposed as a successor to an older theory. The new theory solves some problems the old one could not handle, but there is also some problem for which the old theory had a solution and which the new theory does not solve. Suppose now that the agent accepts an evaluative standard according to which it is rational to accept a successor theory only if it solves all the problems solved by its predecessors. Let us also suppose that he has heard arguments that point to difficulties that would follow if changes involving loss of problem-solutions were allowed, that he is not familiar with arguments against this standard (such as those presented in Laudan 1977), and that this standard is also accepted by his colleagues in the discipline. But suppose further that he, and his colleagues, are in fact mistaken and that this particular standard is not adequate: the ability to solve all problems solved by the predecessor theory is in fact not a necessary requirement for the successor being a better theory.

In refusing to accept the proposed new theory, the agent is doing his rationality duty. He is using the information available to him and applying the evaluative standards he believes to be adequate. He does what he believes to be rational, and to deviate from that would be irrational. From his own, first-person, perspective, this is the best he can possibly do — if rationality is to be action-guiding it cannot require more than this of him. So in acting standard-relatively rational, the agent is doing his rationality duty. Even though he might be acting on the basis of false information or inadequate standards, he is not to be blamed. Given his present state of partial ignorance, he is doing what he rationally ought to.

Still, it might be advisable that he act differently. I am now thinking only of what is advisable in relation to the adequacy of standards (not, for example, in relation to the truth or verisimilitude of theories). Someone who knows that the standard in question is inadequate, could judge that the agent would be better off accepting the successor theory: he would then accept what is in fact, according to the adequate standards, the better theory of the two.

From an external point of view an observer might then offer an advice about, for instance, how to choose between competing theories. Actually receiving such advice changes the situation of the agent. From his own, first-person perspective it might indeed be rational to accept the advice
and act accordingly (if he has reasons to believe the advisor is better qualified to judge the matter and that he ought to be trusted). But for the agent the question always remains: is it rational to accept this advice? And in answering this question, he has no option but to appeal to some standards he himself takes to be adequate.

So an agent might be doing what he rationally ought to, his rationality duty, although he happens to be mistaken about what standards of rational inquiry are adequate. But in trying to act rationally, an agent aspires to act according to standards that are not merely his standards but that are also adequate. To the extent that his standards are inadequate, he will therefore fail in an important sense to do what he himself aspires to do. So although an agent rationally ought to conform to the standards he himself accepts as adequate and thus do his rationality duty, this is not an end in itself. What is of real importance is to find out if there are inadequacies in the standards he has hitherto accepted (or if there are any new, adequate standards that could be added to his present set of standards and that would enable him to conduct his inquiry more efficiently). It is in the agent’s interest that his standards be adequate and that, if he presently has inadequate standards, to detect this and revise his standards (just as it is in his interest to discover new standards that allow him to inquire in new and better ways than before).

Still, in trying to criticize his own standards to detect possible inadequacies, and in judging whether one set of standards is more adequate than another one, an agent must try to be standard-relatively rational. This places some restrictions on the ways in which standards of rational inquiry can be rationally criticized and changed.

### 6.7 Standard-relative and canonical evaluation

In this section I shall discuss two different ways of evaluating the rationality of an inquiring agent’s cognitive actions. These kinds of evaluation I shall refer to as "canonical evaluation" and "standard-relative evaluation".

I am now not discussing the evaluation of theories, arguments and so forth. Instead I shall focus on evaluating the rationality of such evaluations and other such cognitive actions. An agent A may use an evaluative standard S for comparing two theories T₁ and T₂ and deciding
which one of them is best. That is an evaluation of $T_1$ and $T_2$, and the outcome might be for example that $T_1$ is preferable to $T_2$. Using $S$ and arriving at this preference is a cognitive action of A’s. Now we can also evaluate the rationality of A’s action (which led him to this preference), and we may ask questions such as "Is it rational of A to use $S$ for comparing $T_1$ and $T_2$?". If we think that $S$ is adequate and relevant for this purpose, we can ask whether A’s judgment really conforms to the standard $S$ he purports to be using. It may also be that we regard $S$ as an inadequate standard, and judge that according to a different standard $S_1$ which we take to be adequate, A’s evaluation is mistaken and not fully rational.

Usually, when we evaluate an action in terms of rationality, we do so by taking a standard we believe to be adequate (and relevant to the case at hand) and then seeing if the agent’s action conforms to or violates that standard. What usually interests us, that is, is whether an action is an expression of absolute rationality which conforms to standards that really are adequate (we may of course be mistaken about what absolute rationality amounts to and what standards are adequate). In such cases we appeal to the standards that we ourselves, the evaluators, take to be adequate. This kind of evaluation I propose to call a "canonical evaluation".

Since I have already claimed that standard-relative rationality is a genuinely normative notion, it should come as no surprise to the reader that there is also such a thing as standard-relative evaluation. As the term suggests, this is a matter of whether the agent whose actions are being evaluated manages to conform to the standards he himself accepts. Once again, I am not thinking of the triviality that an agent can be described as conforming to or violating his own standards. Since any agent rationally ought to conform to those standards he takes to be adequate, this is a kind of normative evaluation. One evaluates whether or not the agent is doing his subjective rationality duty.

Both these kinds of evaluation are primarily third-person evaluations: an evaluator looks at another agent and decides whether an action conforms, either to standards and rules the evaluator takes to be adequate, or to those the agent being evaluated himself accepts.

An agent contemplating alternative courses of action can of course also evaluate himself in this way. Then standard-relative and canonical
evaluation coincide. But he may also retrospectively evaluate his own past actions (in the third-person way). If he has changed his standards since then, he may for example evaluate a past action as standard-relatively but not canonically rational.

In a canonical evaluation, the standards that the evaluator appeals to are by him taken to be adequate. Actions on the basis of adequate standards are absolutely rational, but I prefer not to call this kind of evaluation "absolute evaluation". The reason is that evaluations are always made with reference to some particular standard or set of standards, but absolute rationality does not stand for any particular standards.

Absolute rationality is not itself to be identified with any particular standards. It stands for the idea of action conforming to perfectly adequate standards, but it is not itself a standard and it is not itself to be identified with any particular standards. Concerning any particular standard one can intelligibly ask whether it is adequate, whether acting according to it is really absolutely rational. Absolute rationality cannot itself function as a standard. It is rather an ideal that we aspire to when we act according to normative standards that we believe to be adequate. Therefore I have chosen a different term, "canonical evaluation".

6.8 Sankey on diversity, change and relativism
In Rationality, Relativism and Incommensurability (1997) Howard Sankey proposes a view about the relation between normativity and adequacy which differs from the one I offer in this chapter. This is interesting, since on the whole Sankey argues for a view of scientific rationality as open to change and conflict, which largely agrees with the general view in my work. In his book, which is about scientific rationality, he argues that standards change over time and that at any given time different scientists may accept different sets of such standards. Standards may according to Sankey change rationally and be improved as science develops, and there are conflicts over what standards should be used in science (Sankey calls his position "methodological pluralism"). So his account of scientific rationality allows for change and diversity with regard to evaluative standards and methodological rules.

Then Sankey asks himself whether this leads to relativism — or, more precisely, to rationality relativism. Rationality relativism as Sankey uses it
is the thesis "that rational theory acceptance is relative to operative standards." (Sankey 1997, 151) By saying that a standard is "operative" he means that it is accepted and used by an agent or a group. I think that this formulation of relativism is too simplistic, but I shall go along with it for a moment. Here is how Sankey expresses the worry that his pluralism leads to relativism:

For if there have been changes in accepted methodology throughout the history of science, then the standards which have been operative in one historical context may differ from those operative in another such context. But this suggests that what it is rational for a scientist to believe varies with context because there has been variation of standards relative to context. And this in turn suggests that scientific rationality is relative to contextually variant standards. Thus, the thesis of methodological pluralism seems to collapse into rationality relativism. (Sankey 1997, 155)

Sankey argues that it does not, that the claim that there is change and diversity of standards does not imply relativism. I agree that it does not, but I think Sankey’s explanation of the matter is unsatisfactory. Since I believe that Sankey articulates an idea which is common among philosophers but which I think is mistaken, I shall examine his argument.

The question, according to Sankey, is "whether variation of operative methodological standards really does imply that rationality varies relative to such standards.” (Sankey 1997, 155) It does not, he claims. He says that the relativist thesis would follow only on the assumption that conformity with operative standards is a sufficient condition for rationality (rational belief). But to make that assumption would be mistaken for two related reasons. First, Sankey says, one must distinguish between "descriptive and normative considerations" (Sankey 1997, 156). To say that a standard is accepted by some agent (operative) and that something conforms to that standard, is a descriptive matter. To say that something is rational is to make a normative claim which goes beyond describing it. This is something I have insisted on, too. Second, Sankey emphasizes, as I have done, that standards may be inadequate. Although a certain standard is operative, there may be "a deficiency in the standard” (Sankey 1997,
138). Thus one can always raise the question whether a given standard is adequate or not, "whether a given standard does indeed provide a basis for rational belief". Sankey’s opinion is that the adequacy of a standard requires that it "yield rational support" (Sankey 1997, 156). He is of course correct in maintaining that standards may fail to do so.

Sankey’s idea is then that if a standard is not adequate then conformity with it is not normatively rational. That is, he takes the adequacy of the standards used to be a necessary condition for rationality. If the standards adopted by an agent or a group are not adequate, then acting (believing) on the basis of them is not rational. Therefore, the change and diversity of standards and rules does not mean that what is rational varies with what standards are operative. So methodological pluralism does not lead to rationality relativism: "Given this potential gap between operative standards and normative rationality, mere adherence to operative standards does not necessarily provide beliefs with rational justification." (Sankey 1997, 157)

This is a way of avoiding reducing normative rationality to conformity with the standards and rules different agents accept. But it is problematic. The trouble is that by making the adequacy of standards a necessary condition for normative rationality, Sankey is left with a form of rationality which is no longer action-guiding. Sankey agrees with me that in the course of inquiry agents may come to discover that standards they have been using are inadequate or deficient. Sankey’s view seems to entail that before such a discovery of inadequacy, the agents in question were not rational in any normative sense at all, although in conforming to the inadequate standard they did the best they can to be rational.

More tellingly, perhaps, the same consideration applies to each of us here and now. Sankey would agree with me that the possibility cannot be ruled out that some standard or standards among the ones we presently accept is in fact inadequate and in need of revision. Let us assume that this is indeed the case, but that we do not now know which standard is inadequate or how it should be revised. What are we to do? That is, how ought we to act? We cannot act on the basis of information that we do not have (information about the inadequacy of one of our standards). The best we can do in this situation, and the only rational thing to do, is to act as we believe to be rational, by conforming to those standards and rules we do believe to be adequate. If one makes the adequacy of standards a
necessary condition for rationality, this would not be rational (since we would be acting on the basis of a standard which, unbeknownst to us, is not adequate). Therefore I think Sankey’s view of the relation between normative rationality and the adequacy of standards is mistaken.

In his attempt to avoid a relativistic conclusion, Sankey disregards the first-person, action-guiding sense of rationality. He has himself stressed that rationality must be viewed as something that is available to inquiring agents (unlike a transcendent goal such as truth). But in making the adequacy of standards a necessary condition for conformity to them being normatively rational, he is in effect sacrificing the action-guiding aspect of rationality. When Sankey makes this move he is left with a kind of objectivistic concept of rationality which has no clear connection with inquiring agents’ resources for cognitive action.

What leads Sankey astray at this point is, I think, an ambiguity in terms like ”rational belief”, ”rational support” and ”rationally justified”, an ambiguity which is due to the association that is often made between rationality and justification and between justification and something like likelihood. To say that a belief is ”rationally justified” may be a normative appraisal to the effect that one ought to accept it, or is permitted to accept it. On the other hand, epistemologists usually assume (or hope) that there is a relation between justification and truth. As Susan Haack puts it, ”truth-indicative is what criteria of justification need to be to be good” (Haack 1993, 203). Thus someone who judges that a belief is justified by meeting the relevant criteria of justification, may think that this belief is in an objective sense supported by evidence and likely to be true (since the criteria are believed to be truth-indicative). But, as Haack emphasizes, whether a belief is justified, and whether justified beliefs are indeed likely to be true, are two separate questions. My hypothesis is that Sankey, in talking about beliefs being rationally justified and rationally supported, mixes these questions up.

There is indeed a gap between ”operative standards” and adequate standards — between standard-relative rationality and absolute rationality. But the action-guiding sense of rationality, which is the one that is relevant to how inquiring agents rationally ought to act in the

---

7 Sankey writes that ”unlike truth in the metaphysical realist sense, rationality is a state which may be determined — albeit fallibly — to obtain or not to obtain from within the perspective of our epistemic practices.” (Sankey 1997, 161)
situations they find themselves in, is standard-relative rationality. I recommend biting the bullet and accepting that the action an inquiring agent rationally ought to perform may consist in acting on the basis of inadequate standards.

6.9 Adequacy and absolute rationality as regulative ideas
The idea of adequacy should be seen as a regulative idea (in Popper’s sense). When we engage in inquiries and make use of different kinds of standards, we aim or aspire to use standards that are truly adequate. Since we cannot ultimately justify standards, we can never be absolutely certain that this aspiration is fulfilled, and the revisability policy advises us to try to detect inadequacies in our standards and to try to make them more adequate. So to have standards that are truly adequate is something that we always aspire to, without ever being able to tell with certainty whether we succeed.

Since absolute rationality is defined as action in conformance to adequate standards, absolute rationality is also a regulative idea. It is something we aspire to as we try to act rationally in our inquiries, although it always remains possible that we fail to achieve it, without our knowing so.

In an addendum to the second volume of *The Open Society and Its Enemies* (Popper 1962), Popper insists on distinguishing between normative proposals and non-normative propositions, and between standards and facts. Concerning propositions, Popper writes that we can always seek truth, and perhaps even find it, but that we can never be certain that we have found it. Absolute truth, in the sense of correspondence with the facts, is to be understood as a regulative idea when it is propositions we are dealing with.

He then adds that the situation with regard to normative proposals and standards is similar. He writes that

> there is some kind of regulative idea about both. In the realm of facts it is the idea of correspondence between a statement or proposition and a fact; that is to say, the idea of truth. In the realm of standards, or of proposals, the regulative idea may be
described in many ways, and called by many terms, for example, by the terms 'right' or 'good'. (Popper 1962, 385)

From the examples he mentions, it seems that when writing about proposals and standards, Popper is primarily thinking of morality; as examples of standards he cites the golden rule and that cruelty is always bad and should be avoided. But what he says about standards and proposals also applies to standards of rational inquiry and the rationality of cognitive action. While Popper writes about standards being good, I have decided to use the term "adequate", and instead of writing about the rightness of proposals, I prefer to speak of the absolute rationality of actions.

Whereas Popper regards the regulative idea of truth as correspondence as clear, he admits that the parallel ideas of goodness (adequacy) and rightness (absolute rationality) are less clear: "It must therefore be admitted that the logical situation of the regulative ideas, of 'right', say, or 'good', is far less clear than that of the idea of correspondence to the facts." (Popper 1962, 385) Nevertheless, he adds that

just as we may seek for absolutely true propositions in the realm of facts or at least for propositions which come nearer to the truth, so we may seek for absolute right or valid proposals in the realm of standards — or at least for better, or more valid, proposals. (Popper 1962, 385-6)

I agree that it is not very informative to say that the adequacy of standards and the absolute rationality of cognitive actions are regulative ideas that we seek to realize when we adopt standards and use them to guide our actions. But it is possible to try to say something more specific about what this adequacy of standards consists in. The instrumental view of adequacy is precisely an attempt at such a specification. Although one could leave the ideas of adequacy and absolute rationality quite open, one can try to give more content to these ideas by stating what adequacy consists in. That is, in effect, what Popper himself does concerning "the realm of facts" when he states that the goal of inquiry should be understood to be truth as correspondence.
What the goal of inquiry should be, and whether truth as correspondence is an appropriate regulative idea, is something that can itself be discussed. Any proposal, such as the idea of truth, can in principle be discussed and criticized. This, Popper would admit. The same thing is true concerning any specification of what adequacy, and hence absolute rationality, consists in. In section 6.11 I shall discuss the consequences of this openness and criticizability of even our conceptions of what it is for standards to be adequate.

First, I shall mention some complications that arise concerning the concept of absolute rationality, if the adequacy of standards is held to be a matter of how well they function as means to goals.

6.10 Complications with the instrumental conception of adequacy
In this section, I shall briefly discuss some complications that arise concerning the concepts of adequacy and of absolute rationality, given that one specifies adequacy in terms of being an efficient means to the goals of inquiries. This is what I call the "instrumental conception of adequacy". This instrumental view of the adequacy of standards is itself open to criticism and the possibility of revision and the consequences of this will be discussed in the next section.

On the instrumental conception, the adequacy of standards is a matter of how they are related, as means, to goals of inquiry. But it seems that what cognitive goals are sought may depend on the subject matter one is inquiring into and on what kind of inquiry one is conducting, and that different agents, working in the same field, may also have different goals. Given the existence of a diversity of goals, it seems that a certain standard may be adequate in relation to one goal but inadequate in relation to other goals. That is, whether a standard is adequate or not, depends on which goal it is taken to be a means to. As an example, let us take the standards of Popper's falsificationist methodology (as presented in Popper 1980). Let us call this set of standards PFM. Next let us look at two different goals that could be proposed for scientific inquiry: (1) the goal of the growth of conjectural knowledge (along the lines of Popper 1989, ch. 10); and (2) the goal of certain knowledge. Let us assume that PFM is an adequate means for achieving (1). On the other hand, PFM is clearly not
an adequate means for someone who wants to achieve goal (2). In this case, PFM is adequate in relation to goal (1), and actions in conformance to PFM are absolutely rational if the agent in question has (1) as his goal. But PFM is not adequate in relation to goal (2), and someone who has that goal fails to be absolutely rational if he conforms to PFM.

No standard is then adequate as such, or in itself, since adequacy is always relative to some goal. Consequently, no action is absolutely rational as such, but only in relation to some goal. But is this really problematic, or is it just an acceptable and trivial consequence of viewing rationality as goal-directed?

An obvious cause of concern, is that this would be a too liberal conception of rationality and of adequacy, in that the actions of agents who have silly or unworthy cognitive goals could qualify as absolutely rational.

A possible way of dealing with this worry, is to introduce an idea of the appropriateness of goals, and say that an action is absolutely rational only if it aims at achieving an appropriate goal and if it is in conformance with standards that are adequate in relation to that goal. This is perhaps a good idea, but what it amounts to depends, of course, on how the concept of appropriateness is specified.

One possibility is immediately suggested by the instrumental conception of rationality and the discussion of Laudan’s and Briskman’s accounts of how to criticize goals rationally (see chapter 5). Their views about how to be rational with regard to goals of inquiry come from within a broadly instrumental perspective: goals may be impossible or difficult to achieve, one goal may fail to guide inquiry properly, goals that are supposed to be means toward more basic goals may be inadequate means, and so on. The kinds of goal criticisms they propose are all of a (broadly) instrumental kind. One possibility is then to say that a goal is appropriate as long as it is not ruled out by this kind of instrumental considerations.

Still, it may be that the constraints that such instrumental considerations impose on the appropriateness of goals, are so weak that a wide variety of cognitive goals would all count as appropriate.

---

8 This would be what Stenmark calls a "holistic" conception of rationality. He formulates the holistic conception as follows: "Rationality consists in the pursuit of appropriate ends and of appropriate means to achieve those ends." (Stenmark 1995, 35)
Perhaps one can make stronger demands on goals, insisting, for example, that to be appropriate a goal must be of real \textit{value} in some objective sense? A related question is if there is some form of \textit{pure value rationality}, which deals with the appropriateness of goals also in non-instrumental terms.

I shall not attempt to answer these questions. It seems reasonable to include in the definition of an absolutely rational action, that the agent is acting to achieve a goal which is appropriate, and that he is conforming to standards that are adequate in relation to that goal. Minimally, I suggest that an appropriate goal should be taken to be one that is not ruled out by instrumental considerations. But I will leave open the question of whether anything more should be required for a goal to count as appropriate.

\section*{6.11 The critizability of conceptions of adequacy}

In this work I have defended the revisability policy, according to which we should hold our standards open to rational criticism and revision. We are not forced, I have argued, to treat any of our standards as in principle exempt from criticism and revision. When we try to criticize a standard we try to find out whether it is adequate or not. The theme of this section is that we can also hold our ideas about \textit{what it is for standards to be adequate} open to rational criticism and revision.

We need, I have argued, some idea of the adequacy of standards, the idea that standards can be good, correct, right, etc. Due to the normativity of rationality, and of standards of rational inquiry, one can always intelligibly ask concerning any standard if it really is adequate. Let us call this the "bare idea of adequacy". The bare idea of adequacy is important, if the concept of rationality is not to collapse into rationality relativism. It also serves to remind us that the normative ideas of adequacy and of absolute rationality should never be equated with what we currently believe about what standards are adequate and about what it is to be adequate. If someone claims, for example, that a certain standard is an optimal means to our cognitive goals, it is still intelligible to ask: "But is it really an \textit{adequate} standard?" The bare idea of adequacy allows us to ask such a question, but it does not say what adequacy consists in, what it is for a standard to be adequate.
One can have different ideas about what the adequacy of standards of rational inquiry consists in. As I have said repeatedly, my own view is that adequacy should be understood in terms of being an appropriate means for achieving the goals of inquiries. But I have also argued, earlier in this chapter and in chapter 5, that it is too simpleminded to think that all standards are merely external means to goals. A correct instrumental view must for example make room for the way in which some standards serve to specify goals, and thus are partly constitutive of rather than externally related to them. Precisely how such a broader instrumental conception of adequacy should be spelt out, I do not know. And apart from such difficulties with the instrumental conception itself, there exist other, non-instrumental conceptions of what it is for standards to be adequate (for example, the one proposed by D’Agostino).

It should be noted that two agents who disagree about what the adequacy of standards consists in, do not have to disagree about which standards are adequate. They may share standards which enable them to conduct their debate without begging the question in favour of either alternative. Similarly, if an agent revises his conception of adequacy, it need not be the case that he revises any of his standards. At least some of his standards will remain as a temporarily fixed point which allows him to evaluate and revise conceptions of adequacy.

Particular ideas about what this adequacy consists in, are really competing specifications of the general, bare idea of adequacy. And just as the claim that a certain standard is adequate, the claim that adequacy consists in one thing rather than another is in a sense hypothetical. It is a claim that can intelligibly be challenged, and just like standards and goals, conceptions of what adequacy of standards consists in can be revised and improved. Even this, what we take rationality in inquiry to be in the first place, is something we can treat as an open question.

6.12 Relativity without relativism
By way of conclusion, I want to make clear that the account of the normativity of rationality in inquiry that I have presented in this chapter does not amount to rationality relativism. Please note that the following is not an argument against relativism. I am merely trying to show that what I have said in this and in earlier chapters does not have rationality
relativism as a consequence, and that instead it is inconsistent with such a relativism.

To make it clearer what this relativism that I claim to avoid is, I shall begin by presenting a couple of related definitions made by others.

First there is Hartry Field, who writes that by "epistemological relativism" he means

the doctrine that the basic epistemological properties are not such properties as that of belief $B$ being justified, but rather such relativized properties as that of belief $B$ being justified relative to evidential system $E$. (An evidential system is, roughly, a bunch of rules for determining under what conditions one is to believe various things; a belief is justified relative to an evidential system in certain circumstances if the rules license the belief under those circumstances.) (Field 1982, 563)

Instead of "belief" I will talk about "cognitive action", and I shall use "rational" and "set of standards" instead of "justified" and "evidential system".

The central point about this epistemological relativism sketched by Field is that there is no unrelativized, absolute sense of rationality (justification), but that there is only rationality relative to the set of standards of different agents (justification relative to evidential systems embodied in people’s epistemic practices). The conceptual "resources" of Field’s relativist "don’t allow him to say that a belief is justified relative to the true evidential system", the reason being that "since the last use of the term 'justified' is unrelativized, the epistemological relativist cannot say this." (Field 1982, 563) Translated into the terminology I use here, the rationality relativist would say that there is no such thing as absolute rationality, no such thing as an action being rational by standards that are adequate — since a standard cannot be adequate as such over and above being accepted, or embodied in a practice. Conversely, there would be no sense in which an agent’s standards could be objectively inadequate, either.

Another thing that I want to pack into the concept of rationality relativism is brought out in a characterization of "evaluative relativism"
by Rom Harré and Michael Krausz. This relativism says not only that people differ "with respect to the standards by which they evaluate their beliefs", but also "that there could be no overarching set of evaluative standards by which each person's evaluative standards could be ranked with respect to its knowledge engendering power" (Harré and Krausz 1996, 74). I take this to mean that rival standards or sets of standards cannot be rationally compared, and that there is no way of rationally criticizing an agent's set of standards.

Let me put these pieces together into a list of theses. Together they make up what I shall mean by "rationality relativism".

According to rationality relativism:

1. What cognitive actions are rational is relative to the standards accepted by agents.
2. Standards cannot be adequate, or inadequate, as such.
3. Therefore there can be no such thing as absolute rationality;
4. there is only rationality relative to the standards accepted by different agents.
5. When two agents $A_1$ and $A_2$ evaluate a theory $T$ by different standards $S_1$ and $S_2$, the results of their evaluations cannot conflict with each other: what looks like a conflict really just consists in (e.g.) $T$ being rational-for-$A_1$-by-$S_1$ and $T$ not being rational-for-$A_2$-by-$S_2$ — and these things are not incompatible with each other.
6. The standards an agent accepts cannot be criticized.
7. Different sets of standards cannot be rationally compared.

(1) is true according to the view I present in this chapter. What cognitive actions an agent rationally ought to perform (what his subjective rationality duty consists in) does depend on the set of standards accepted by the agent in question. An agent rationally ought to try to conform to those standards he takes to be adequate.

(2), (3) and (4) are false if my account is correct. Standards can be adequate and inadequate, independently of whether agents accept them or not. There is, contrary to (3), such a thing as absolute rationality: action in conformance to standards that are fully adequate. Hence (4) is false,
too, because besides the concept of standard-relative rationality there is the concept of absolute rationality.

Absolute rationality is what inquiring agents aspire to as they try to conform to their own standards: they accept them as theirs and use them to guide their inquiries because they take them to be adequate standards. Therefore agents such as A1 and A2 mentioned in (5) do disagree: between them there is a real conflict concerning which standard is adequate for the purpose of evaluating T and thus concerning which action (evaluation) is absolutely rational.

Neither (6) nor (7) follows from the view presented here. I have said that standards may be inadequate, and that it may be perfectly legitimate to evaluate an agent’s action as canonically irrational because it is based on standards one takes to be inadequate. In this chapter I have not discussed the issue of how standards can be rationally criticized and how sets of standards can be compared. However, in chapters 3 to 5 I have argued that there are no good reasons for treating any standards as uncriticizable, and I have given examples of ways in which standards can be criticized and compared. So if what I say in earlier chapters is correct, it follows that (6) and (7) are false.

It might still be objected that I have conceded too much to the relativist, and that the claims I have made about the standard-relativity of rationality can be used to argue in favour of rationality relativism. I think that precisely the opposite is the case. The account of rationality in inquiry which I have presented in this work — claiming among other things that we are fallible with regard to standards, that all standards should be treated as revisable and that it can be normatively rational to conform to inadequate standards — actually weakens the case for relativism. They show how an absolutist (non-relativist) conception of rationality in inquiry is compatible with those observations about rationality and standards that have been appealed to in arguments for relativism: that we cannot prove that standards are adequate, that different agents accept different standards, and that standards change over time. All of this can be acknowledged by an absolutist rationalist, if the relation between the adequacy of standards and the rationality of cognitive actions is properly understood. If the account I have presented is correct, our fallibility and the change and diversity of standards of rational inquiry do not support relativism.
According to the absolutist view of this book, it is indeed true that what it is rational for inquiring agents to do depends on what standards and rules they accept. Thus there is some truth in rationality relativism — which is nevertheless absolutely false.
REFERENCES


The series UMEÅ STUDIES IN PHILOSOPHY consists mainly in research communications (monographs, dissertations and collections of essays) in Philosophy from the Department of Philosophy and Linguistics at Umeå University. The series is not distributed through the bookstores, but individual volumes can be obtained from the department. Orders should be sent to Umeå Studies in Philosophy, Department of Philosophy and Linguistics, Umeå University, SE-901 87 Umeå, Sweden.
