

THE INSTRUMENTAL CONCEPTION OF EPISTEMIC RATIONALITY

A PHILOSOPHICAL EXAMINATION OF THE INSTRUMENTAL CONCEPTION OF  
THE EPISTEMIC RATIONALITY OF HUMAN DOXASTIC STATES

By PATRICK BONDY, B.A.[H], M.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the  
Requirements for the Degree of Doctor of Philosophy

McMaster University

© Copyright by Patrick Bondy, August 2012

DOCTOR OF PHILOSOPHY (2012) (Philosophy)

McMaster University

Hamilton, Ontario

TITLE: A Philosophical Examination of the Instrumental Conception of the Epistemic  
Rationality of Human Doxastic States

AUTHOR: Patrick Bondy, B.A.[H], University of Windsor, M.A., University of Windsor

SUPERVISOR: Professor D. Hitchcock

NUMBER OF PAGES: vi, 233

## ABSTRACT

The instrumental conception of epistemic rationality is the view according to which beliefs, or doxastic states generally, are epistemically rational insofar as they promote the achievement of an epistemic goal, and they are epistemically irrational to the extent that they fail to promote such a goal. The thesis that I defend here is that the instrumental conception is not satisfactory as a general account of epistemic rationality.

I proceed by examining a number of reasons one might offer for accepting the instrumental account, and I find them wanting. I also consider various ways of formulating the epistemic goal, attempting to determine the best one, in order to show the instrumental conception in its best light. I consider and reject the attempt to ground the instrumental conception on the proper function of our cognitive systems. Finally, I consider three arguments against the instrumental conception of epistemic rationality, and some objections to them. I conclude that, even shown in its most favourable light, the instrumental conception cannot give us a satisfactory general account of epistemic rationality.

I would like to express my gratitude to my supervisor, Professor David Hitchcock, for all of his helpful feedback and encouragement throughout the development of this dissertation. I would also like to thank Professors Mark Vorobej and Brian Garrett for their thorough comments and criticisms. Finally, I would like to thank Kari, for putting up with me this whole time.

I also gratefully acknowledge that this research was supported by the Social Sciences and Humanities Research Council of Canada.

## TABLE OF CONTENTS

Chapter 1: Introduction	1
1.1. Some Groundwork	4
1.2. A Preview	30
<u>PART I: Exploring the Instrumental Conception</u>	
Chapter 2: Instrumentalists and their Critics	36
2.1. Critics of Instrumentalism	36
2.2. Defenders of Instrumentalism	39
2.3. Addendum	60
2.4. Summary	62
Chapter 3: Reasons for the Instrumental Conception	64
3.1. A Deep Explanation	66
3.2. Naturalism	70
3.3. Reasons-Internalism	79
3.4. The Default	84
3.5. A Unified Account of Rationality	86
Chapter 4: The Epistemic Goal	89
4.1. Some Preliminaries	89
4.2. Epistemic Goals and Values	93
4.3. Matters of Interest and Importance	108
4.4. Synchronic vs. Diachronic	117
4.5. Summary: The Epistemic Goal	123

<u>PART II: Objections to the Instrumental Conception</u>	
Chapter 5: Proper Functions and Epistemic Rationality	125
5.1. The Concept of a Function	126
5.2. Proper Functions and Cognitive Systems	131
5.3. Proper Functions and Normativity	144
Chapter 6: Instrumental Rationality is Not Sufficient for Epistemic Rationality	149
6.1. The Problem	152
6.2. Objection 1: No Doxastic Voluntarism	155
6.3. Objection 2: Rule-Consequentialism	157
6.4. Objection 3: The Myopic Restriction	158
6.5. Sufficient Reflection and Properly Based Beliefs	164
Chapter 7: The Value of the Epistemic Goal	176
7.1. Movie-Goers, Unfaithful Spouses, and Grains of Sand	176
7.2. Some Objections	182
Chapter 8: A Regress Argument Against the Instrumental Conception	202
8.1. A Restriction on Instrumental Rationality	202
8.2. The Regress	206
8.3. Three Objections	207
Chapter 9: Conclusion	215
References	219

## Chapter 1: Introduction

Truth acquisition is often desired and enjoyed for its own sake, not for ulterior ends. It would hardly be surprising, then, that intellectual norms should incorporate true belief as an autonomous value, quite apart from its possible contribution to biological or practical ends. (Goldman 1986, p. 98)

Other things being equal, it is good to believe what is true and only what is true. (Lynch 2004, p. 47)

All men by nature desire understanding. (Aristotle, *Metaphysics* 980a21)

Claims like these have a truistic feel to them. Reflection on such truisms is apt to lead us to incorporate true belief into our epistemology, as a goal or end to be promoted. Indeed, although I do not favour instrumentalism about epistemic rationality, I am happy to grant that true belief is often good, and that we should give truth a central role to play in epistemology. I do not mean to oppose all talk of goals and truth-achievement in epistemological theorizing.

The purpose of this dissertation is to inquire into what it is for a doxastic attitude, such as belief, to be epistemically rational.<sup>1</sup> In the literature of the past few decades, epistemic rationality of this sort has often been conceived to be instrumental in nature. According to the instrumental conception, epistemic rationality has to do with holding beliefs that are appropriately related to an epistemic goal. On this view, a belief is epistemically rational if and only if holding it promotes the achievement of some specified epistemic goal, or it is reasonable to think that it promotes the achievement of

---

<sup>1</sup> Other things might be thought of as epistemically rational or irrational (methods of inquiry, for example). Such things are not the subject of this inquiry. Claims I make about how to characterize epistemic rationality and irrationality should be understood only to apply to beliefs and other doxastic states, where I understand doxastic states to include belief, disbelief, and suspension of belief. Degrees of belief, if there are such things, are also a type of doxastic state.

the epistemic goal, or it meets some other condition along those lines. The most common way to think of the epistemic goal, following William James (1949/1896), involves achieving truths and avoiding errors. We will see later that just how to formulate the epistemic goal is a matter of some debate, but almost everyone agrees that it has something to do with having true beliefs and avoiding false ones.

The thesis that I shall defend here is that epistemic rationality is not best conceived as instrumental in nature. The instrumental conception<sup>2</sup> does have some appealing features, but it also faces some serious obstacles, and I argue that those obstacles warrant abandoning the view that epistemic rationality has to do with the achievement of a goal.

I should make it clear, however, that I do not mean to argue that there are no interesting instrumental evaluations of beliefs and belief-forming practices to be made. Sometimes people want to know the truth, and sometimes it is good for us to know the truth, even if we'd rather not. Sometimes these kinds of evaluations might be relevant for epistemology. But the point of this thesis is that a thoroughgoing instrumental approach cannot deliver an adequate account of epistemic rationality.

Throughout this dissertation, I will frequently appeal to either cases where subjects hold beliefs against the evidence, as clear cases where we would judge the

---

<sup>2</sup> In many places, I will refer to the instrumental conception of epistemic rationality as “the instrumental conception,” “the instrumental approach,” and “instrumentalism.” The use of these variations is only a matter of style, to avoid repeating “the instrumental conception of epistemic rationality” ad nauseum. All of these expressions are meant to refer to the same thing. Also, there is no clear distinction in the literature between “instrumental” and “teleological” accounts of epistemic rationality. I will predominantly use “instrumental,” but I will sometimes switch to using “teleological,” since some theorists use that term.

beliefs to be epistemically irrational, or cases where subjects hold beliefs based on good evidence, as clear cases where we would judge the beliefs to be epistemically rational. The method of considering common-sense judgments about the applicability of a term, and employing those judgments as defeasible but important constraints on our theorizing, is an important and widely employed method of philosophical analysis, and it is a stock method of post-Gettier epistemology, as well as of many important figures throughout the history of philosophy. Of course, consulting our intuitions about cases and concepts is not a very good way to learn new things about the world. However, it *is* a good way to learn about our concepts, and what I am interested in is the concept of epistemic rationality – just what does it mean to say that a belief is epistemically rational? To what kinds of cases does the concept apply? The responses to those questions which are embodied in the instrumental conception of epistemic rationality essentially involve reference to the achievement or promotion of some relevant epistemic goal(s). The purpose of this dissertation is to explore that response, and to argue that it is ultimately unsatisfactory.

As we shall see, prominent instrumentalists about epistemic rationality are also concerned to respect the importance of evidence in determinations of epistemic rationality. The method of appealing to cases where there is good evidence at hand, and determining whether the instrumental approach can yield the correct verdict in such cases, therefore does not beg the question against the instrumental approach.

This chapter has two parts. In the first, I set out some important conceptual groundwork, setting out the distinction between epistemic and practical rationality, the

instrumental and categorical conceptions of epistemic rationality, and the relation between epistemic reasons and epistemic rationality. The second part consists of an overview of the arguments to come in later chapters.

## **1.1. Some Groundwork**

### **1.1.1. Epistemic and practical rationality; rationality and justification**

First of all, it is important to highlight the distinction between epistemic and practical rationality. This distinction is ubiquitous in the literature, but it is not often treated in any great detail, so we would be well advised to dwell on it a moment.

As a first pass, it is tempting to say that practical rationality is the rationality of action, and epistemic rationality is the rationality of belief. That way of drawing the distinction isn't quite right, though, because beliefs can be subjected to both epistemic and practical evaluations. It can, in some cases, be practically rational to hold epistemically irrational beliefs. Consider the following example, drawn from Heil (1992):

#### **Alexei**

Alexei has good evidence to think that his wife, Anna, is being unfaithful. Alexei knows that if he does not believe that Anna is faithful, then he will be unable to prevent himself from behaving coldly towards her. He also knows that, if he behaves coldly towards her, given the poor state of their marriage already, the marriage will end. But he desperately does not want his marriage to end. So, all things considered, it would be best for Alexei to believe that Anna is faithful, and when she tells him that everything is all right and just to trust her, he believes her.

The thrust of this example is that, despite the fact that it is clearly epistemically irrational, Alexei's belief that Anna is faithful to him can be practically rational, because it is the means to achieve a goal that is very important to Alexei.

Here is another example to illustrate the same thing:

### **Winston**

Winston lives in a totalitarian society, in which people get into serious trouble for saying anything to the effect that anyone living in that society lacks personal freedoms. Winston knows that, if he believes that the society is unfree, he is likely to end up saying things to that effect, which will be very dangerous for him. It would be prudent, therefore, for Winston to believe that he really does not live in a society that lacks personal freedoms, even if it would not be epistemically rational.

In Winston's case, we can evaluate his belief that he lives in a free society (if he does bring himself to believe it) as both epistemically irrational and practically rational. So, because beliefs can be practically rational, it is not a distinguishing feature of epistemic rationality that it is the rationality of beliefs.<sup>3</sup>

How can we begin setting out the fundamental idea of epistemic rationality, if it is not exclusively the rationality of beliefs? Well, one point to notice is that there is a trend in recent epistemology to use "epistemically justified belief" and "epistemically rational

---

<sup>3</sup> Mills (1998) does not accept that epistemic and practical rationality can come apart like this, but almost everyone else accepts it. Even Heil, who seeks a unified account of epistemic and practical rationality, allows that the two may come apart in particular cases, despite the fact that he holds that the purpose of epistemic rationality is to further our practical interests. See his (1983) and (1992).

belief” synonymously.<sup>4</sup> If that is correct, then we can get at the basic idea of epistemic rationality by considering the concept of epistemic justification.

Epistemological lore has it that epistemic justification has been traditionally thought of as what turns true belief into knowledge. The idea is that knowledge cannot be mere true belief, because even beliefs that are held for bad reasons can luckily turn out to be true. A very old view (Plato’s), for example, has it that knowledge is true opinion that is appropriately tied down or secured, so that it does not too easily fly away.<sup>5</sup> What has traditionally been thought to be that which secures true belief is justification. And justification does that job, because a belief is epistemically justified if there is reason to think that the belief is true.

There is disagreement among epistemologists over whether the reason for thinking that the belief is true must be accessible to the subject who has the belief, or whether it can be external to the subject’s cognitive perspective. Internalist evidentialists, for example, hold that a subject needs to have cognitive access to evidence that supports the belief in order for it to be epistemically justified; externalist reliabilists hold that a belief is epistemically justified just in case it is produced by a reliable process, whether or not the subject has cognitive access to the process itself. We do not need to concern ourselves with the internalism/externalism debate here. The point is that epistemic justification has to do with reasons for thinking that a belief is true, however we want to conceive of the

---

<sup>4</sup> Cf. Boghossian (2006, p.14n) and Cruz and Pollock (2004, p.125n), who make it explicit that they use the terms interchangeably; many others (e.g. Laudan (1990), Bergmann (2006a)) use them that way implicitly.

<sup>5</sup> See his *Meno* and *Theatetus*.

relation of cognitive access between a believer and the reasons for the beliefs that she holds.

That point remains true even if we stop worrying about how to give an account of justification that turns true belief into knowledge. This is an important point, because Gettier cases<sup>6</sup> are widely taken to show that epistemic justification does not by itself turn true belief into knowledge. Consider the following Gettier-style example:

### **Bad Luck**

Bill believes that Cindy will die this morning, because he saw her fall out of an airplane this morning, without a parachute. But Cindy gets lucky, lands correctly in deep water, and survives, managing to swim to a nearby shore. Then, sadly, before the morning is out, she is hit on the head by a falling coconut and dies.

Bill's belief is very well justified in this case, because the fact that someone falls out of an airplane makes it extremely likely that that person will shortly die. However, that justification failed to turn Bill's belief into knowledge, despite the fact that the belief is true. The justification does not bear on the reason for which the belief is true – it is only by luck that Bill's belief is true. (Good luck for Bill's belief, bad luck for Cindy.)

Some epistemologists think that the Gettier problem can be solved, and that we can give an account of justification that eliminates all such cases, or else that we can add some fourth condition to the analysis of knowledge to eliminate Gettier cases. I do not

---

<sup>6</sup> Gettier cases, so-called after Gettier's seminal (1963) paper, are cases where a subject has a justified true belief, but where the justification for the belief misfires – the justification fails to be properly connected to the truth of the belief. Such cases are counterexamples to the sufficiency of justified true belief for knowledge.

propose to go into the analysis of knowledge here. Richard Foley notes that the two fundamental questions of epistemology are “what is involved in having good reasons to believe a claim,” and “what is involved in meeting the higher standard of knowing that a claim is true?” (2008, p. 42). Although there has traditionally been a tendency to think that the same answer can be given to both questions, Gettier cases show that the relation between the answers to the two questions is not as straightforward as we might initially think (or hope). Foley’s recommendation (in my view a good one) is to relax the tie between these two questions, and to tackle them separately. The study of knowledge and the study of justification, although they are obviously related and have traditionally gone together, ought not to be beholden to one another.

What is important for our purpose is that epistemic justification rules out some but not all cases of epistemic luck, and that it has to do with reasons that have to do with the truth and falsity of beliefs. All of that remains true, even if we no longer worry about whether justification understood in this way is adequate for the purpose of giving an account of knowledge.

One might wonder, at this point, about the status of evaluative beliefs, or beliefs about what will be practically best to do, etc., which might have no truth-value – can beliefs of this kind be epistemically justified? My own view about such beliefs is that such beliefs do have truth-values, and that they can be epistemically rational or irrational. However, I do not need to take a stand on that issue here, for if it turns out that such beliefs do not have truth-values, then it might simply turn out that there is no epistemic

rationality or irrationality involved in holding them, especially for people who know that such beliefs have no truth-value. Perhaps such beliefs could only be practically justified (justified by their fruits, as they say).<sup>7</sup>

My suggestion, then, is that we should treat epistemic rationality as identical with epistemic justification, and that epistemic justification is essentially connected with reasons for thinking that a proposition is true. I should note before moving on that some theorists, like Audi (2001), prefer to keep the concepts of justification and rationality distinct.<sup>8</sup> There are two good reasons for wanting to do so: first, judgments of rationality can be applied to the overall epistemic characters of persons, not only to beliefs, whereas judgments of justification apply only to beliefs; and second, ‘rationality’ as applied to beliefs is ambiguous between its strong and weak senses, where the strong sense is equivalent to justification, and the weak sense indicates merely a lack of positive *irrationality*, such as the conscious holding of explicitly contradictory beliefs. On the weak sense of ‘rationality,’ a belief can be rational provided that there is no positive reason to think that it is false, or when it is simply based on a reason of some kind (even a

---

<sup>7</sup> However, the issue might be more complicated than the way I have presented it here. There might be good reasons for thinking that a practical or evaluative belief is true – even if it can have no truth-value. Suppose someone in a position of epistemic authority tells you that it is *true* that one course of action is better than another. And suppose that practical beliefs cannot be true or false. You might still be epistemically justified in believing what you have been told, if you trust the epistemic authority who says that it is true. But none of that affects the point I am trying to make, which is that a belief is epistemically justified or rational if there is good reason to think that it is true. All it shows is that there can be good reason to think that a belief is true, even if the belief turns out to have no truth-value. That is not so very different from saying that there can be good reason for thinking that a belief is true, even if it turns out to be false.

<sup>8</sup> Goldman (1986) wants to stop using the term “rationality” altogether, due to the fact that it is used in too many senses to be of any use. However, “justification” is used in almost as many ways, and we still (with the exception of Alston (2005)) use that. In any case, many epistemologists have continued talking in terms of rationality, and I shall do so as well.

bad one). So, for example, if a belief is held when there are no reasons either for or against it, it can be rational in the weak sense to hold it. However, because the interest in the theory of epistemic rationality is driven mainly by an interest in the rationality of beliefs, rather than of persons, and it is an interest in rationality in the strong sense, the relevant sense of “epistemic rationality” is synonymous with “epistemic justification.” Rational belief in the strong sense requires more than that there simply be no reason that counts against the belief; there must be something positive counting in favour of the belief. Note, however, that there can be beliefs that are epistemically rational in the strong sense to only a weak *degree*, when there is some consideration that positively favours a belief, but only a little bit. Epistemically rational belief, one might say, is the *contradictory* rather than the *contrary* of epistemically irrational belief;<sup>9</sup> there is no middle ground between epistemic rationality and irrationality, where neither applies. (In cases where there is no evidence either for or against a belief, it might be supposed that the belief is neither epistemically rational nor irrational. It seems to me that holding a belief when there is no evidence either way is epistemically irrational, though.) But bear in mind that beliefs can also enjoy a weak degree of *justification*, as when there is some small presumptive reason to think that the belief is true, and no positive reason to think that it is false.<sup>10</sup> So the fact that beliefs can enjoy a weak degree of rationality does not distinguish rationality from justification. (Just how likely a rational belief must be in light

---

<sup>9</sup> However, throughout this dissertation, the cases we will consider will typically be cases where there is a strong degree of epistemic rationality or irrationality, because such cases are clearer and less controversial than those that are rational to only a weak degree.

<sup>10</sup> Or, say, for a reliabilist, when the belief is produced by a belief-forming process that is only moderately reliable.

of the considerations which count in favour of it is a thorny question that I will not address here, because it is a difficult question that does not need to be settled in order to inquire about the *nature* of epistemic rationality or justification.) In what follows, I will go back and forth between talking in terms of rationality and talking in terms of justification, keeping as much as possible with the usage of the authors in question as we go along.

What we have just been considering are various issues that arise with respect to the characterization of epistemic rationality. It was pointed out that epistemic justification/rationality has essentially to do with truth. Returning now to the relation between epistemic and practical rationality, I think it is best to leave that relation indeterminate for now. Practical rationality is concerned with taking the appropriate means to satisfy our desires or achieve our goals, and perhaps with adopting the right goals, and perhaps also with following certain other practical constraints. I know of no universally accepted way to characterize practical rationality; I intend to remain as neutral as possible here between competing accounts of practical rationality.

What is important about practical rationality, for our purpose here, is that the rationality involved in taking the means to achieve one's goals is at least one species of practical rationality. So, considering again the case of Alexei, it is enough to notice that the goal that Alexei wants very badly to achieve, the preservation of his marriage, can be promoted by his believing that Anna is faithful. Holding the belief against the evidence is

practically rational, then, because holding it is a necessary means to achieve a goal that Alexei cares very much about achieving.

A possible objection to the instrumental conception of epistemic rationality, at this point, would be that it cannot account for the divergence between epistemic and practical rationality. After all, if epistemic rationality is a species of practical rationality, it can hardly diverge from practical rationality. But the instrumentalist about epistemic rationality would account for the divergence between the practical and epistemic justification of Alexei's belief by noting that his holding the belief is rational with respect to a goal that Alexei cares very much about achieving, i.e. the goal of keeping his marriage intact, but it is irrational considered in light of another important goal, the *epistemic* goal. Although holding the belief blocks his achievement of the epistemic goal, and is practically irrational considered only in light of that goal, it is overall practically rational for him to hold the belief, given that he cares more about satisfying the conflicting goal of saving his marriage. Because goals that an agent values more highly trump goals that she values less highly, when her goals come into conflict, it is overall practically rational to take the means to achieve the more highly valued goal.

In other words, if there is in fact an epistemic goal that we want to achieve, or if there is an epistemic goal that we practically *ought* to achieve, and if epistemic rationality is instrumental rationality in the service of the epistemic goal, then epistemic rationality is a species of practical rationality, distinguished from other species by its distinctively

epistemic goal.<sup>11</sup> The instrumentalist can account for Alexei's beliefs by holding that Alexei's belief is practically rational, even though it interferes with achieving the epistemic goal, because it achieves a goal that Alexei cares about more than he cares about the epistemic goal.

On the other hand, if it is not the case that there is an epistemic goal that we either want to achieve or practically ought to achieve, or if epistemic rationality just does not have to do with achieving a goal (i.e. if the instrumental conception is mistaken), then epistemic and practical rationality might simply turn out to be distinct kinds of rationality. I do not want to prejudge the issue at this point either for or against the instrumental conception of epistemic rationality, so for now we can leave it undecided just what is the relation between epistemic and practical rationality. It is enough to notice, first, that epistemic rationality is a status of doxastic attitudes that has to do with reasons to believe that a proposition is true, and second, that instrumental rationality (the kind of rationality involved in taking the means to one's ends) is at least a species of practical rationality.

### **1.1.2. Instrumental and categorical conceptions of epistemic rationality**

A second important distinction is that between instrumental and categorical conceptions of epistemic rationality. Although these terms are appropriate for drawing this distinction, they come with a certain amount of philosophical baggage, so it is important to make their meanings clear. The basic distinction is between conceptions of epistemic rationality

---

<sup>11</sup> That claim rests on the view of instrumental rationality as (at least one type of) practical rationality. Perhaps someone might want to object that not all instrumental rationality is practical, on the grounds that there can be non-practical goals. It seems reasonable to me to think that instrumental rationality is in every case a kind of practical rationality. However, nothing big will depend on that claim in this dissertation; I intend to remain neutral with respect to the relation between epistemic and practical rationality.

as goal-directed (instrumental conceptions) and conceptions of epistemic rationality as independent of the achievement of a goal (categorical conceptions). Epistemologists who take the instrumental approach typically identify a goal (or set of goals) that they take to be characteristically epistemic, and they hold that beliefs that are epistemically rational are those that are held in such a way as to promote the achievement of that goal. (In what follows, I will sometimes say that according to the instrumental conception, epistemic rationality *depends on* the content of an epistemic goal. By that, I mean that the epistemic rationality of a belief is a matter of its being held in such a way as to promote the achievement of the epistemic goal. This way of talking is meant to be neutral between conceptions according to which the belief must *actually* promote the achievement of the epistemic goal, and those according to which the subject in question only needs to have good reason to *think* that it promotes the achievement of the goal, even if it does not in fact promote the achievement of the goal.) We can postpone a detailed discussion of just how the epistemic goal ought to be formulated until the fourth chapter; for now, we can work with Alston's (1985) intuitively plausible formulation, which is to maximize truth and minimize falsity in a good-sized body of beliefs.

Within the instrumental conception, there is a further distinction to be drawn, between conceptions that make epistemic rationality depend on the epistemic goals that agents actually have (care about, want to achieve), and those that make it depend on the content of an epistemic goal that is good to achieve whether or not agents in fact care about achieving it. We can refer to the view that epistemic rationality depends on the

content of the epistemic goals that agents actually have (or care about achieving) as the *subjective* instrumental conception of epistemic rationality. The view that epistemic rationality has to do with achieving an epistemic goal that is good to achieve, independent of whether agents care about achieving it, we can call the *objective* instrumental conception of epistemic rationality. This distinction must be kept carefully distinct from a different distinction that sometimes is cast in subjective/objective terms: the distinction between the means that one *believes* will achieve one's goals, and the means that will *in fact* achieve one's goals. That is a different distinction. When I say "subjective instrumental conception of epistemic rationality," the term "subjective" is intended to pick out the kind of goals that are relevant for the assessment of rationality.

Another way to put this distinction is in terms of the normativity of epistemic reasons. According to the subjective instrumental account, epistemic reasons, reasons having to do with promoting the achievement of an epistemic goal, cannot be normative in the absence of an agent's having adopted that epistemic goal. On this account, epistemic reasons cannot *give* an epistemic agent a reason to believe a proposition, which the agent would be rationally at fault for ignoring, if the agent has not adopted an epistemic goal. On this view, for a subject who does not care about the epistemic goal, there is no epistemic rationality or irrationality. According to the objective instrumental conception, by contrast, it is a rational failing for an epistemic agent to fail to properly take account of epistemic reasons, even if she does not desire to achieve the epistemic goal.

### 1.1.3. What about hypothetical rationality?

One might want to hold that there is another way to talk about instrumental rationality: we simply posit some epistemic goal, whether or not anyone cares about it, and whether or not anyone ought to care about it (i.e. whether it is good or valuable to achieve independent of people's desires), and we make epistemic judgments relative to that goal. Whether the agents in question have any reason to hold beliefs in accord with what is epistemically rational for them depends on whether they care about the epistemic goal, but the evaluation of epistemic rationality itself does not depend on their caring about the epistemic goal, or on whether the epistemic goal would be good to achieve in any sense.<sup>12</sup>

There are two responses to make to the suggestion that this is a viable alternate type of instrumentalism. First, means-ends evaluations of rationality involving ends that are not in fact desired, and which are not goals that ought to be promoted, are really an odd type of evaluation – or at least, it is odd to make such an evaluation and call it an evaluation of rationality. Take the goal of counting all of the hairs on all of the polar bears in the world, for example: surely it makes no sense to posit such an odd goal that no one wants to achieve, and that no one really ought to achieve, and make evaluations of rationality relative to that goal. Evaluations of instrumental *effectiveness*, of course, do not depend on the claim that the goal in question is valuable in any way, but evaluations of instrumental *rationality* have to connect up somehow with the ends that people either do in fact have or that it would be good for them to have. More generally, if an action or

---

<sup>12</sup> James Maffie (1990) and Roger White (2007), for example, hold this view.

belief is to be instrumentally valuable, in the service of some goal, it must be the case that that goal is valuable to achieve in some sense – either because it is in fact desired, or because it is good to achieve independent of whether anyone desires it.

Consider the following example in support of that general claim.

### **World War**

Moriarty intends to begin a world war, so that the arms manufacturing companies that he owns will make him a fortune. He proceeds to take the means to begin to achieve that goal: he arranges the assassination of an Austrian archduke.

Holmes wants to stop Moriarty. He proceeds to calmly consider Moriarty's actions, determining whether they appear to be instrumentally rational. Holmes reflects: "The political situation has indeed become dangerously unstable since the archduke's assassination. Having the archduke killed appears to have been an effective way to achieve Moriarty's goal of starting a world war. Indeed, it was eminently instrumentally rational."

Watson, having lately been engrossed in a series of apocalyptic stories, exclaims: "But my dear Holmes! Having the archduke killed was also a very good first step toward the goal of annihilating life on Earth!"

Holmes impatiently replies, "Come, come, Watson. We can't say that his action is rational insofar as it promotes *that* goal – for he does not, after all, want to extinguish life on Earth."

It seems to me that Holmes's response here is perfectly reasonable. If Moriarty does not want to extinguish life on Earth, and if (as is no doubt the case) extinguishing life on Earth is not a good thing to achieve, then his action is not instrumentally rational insofar as it promotes that goal. At most, we could say that Moriarty's action of having the archduke killed *would* have been instrumentally rational in the service of that goal, *if* he had wanted to achieve it (or if it had been good to achieve it). This example supports the claim that evaluations of instrumental rationality depend on taking the goal in light of which the evaluation is made to be valuable in some sense.

An instrumentalist might want to object, however, that it does not matter whether particular individual epistemic agents care about achieving the epistemic goal, or think that it is valuable to achieve; the fact is that most of us do think that finding out the truth is generally valuable, and that is enough to get the practice of evaluating beliefs in light of the epistemic goal off the ground. So, since enough people do care about achieving true beliefs most of the time, it makes sense to posit an epistemic goal and evaluate people's beliefs in light of it, even if some of those people do not care about achieving that goal.

This objection is partly correct. If enough of us do think that a goal is worth achieving, then we can evaluate people's actions or beliefs in light of that goal, even if those being evaluated do not think that the goal is worth achieving. But notice that positing a goal in light of which to evaluate people's actions or beliefs, when the agents being evaluated do not care about the goal in question, only makes sense because we, the evaluators, take the goal to be valuable. But that just means that we are making an

objective instrumental evaluation of rationality, in the sense specified above, where the goal is taken to be valuable to achieve whether or not the agent whose beliefs we are evaluating cares about it. So we are still only evaluating people's beliefs as instrumentally epistemically rational, in the service of some appropriately truth-centered epistemic goal, when we either take them to care about achieving the goal, or we take the goal to be good to achieve whether or not they care about it.

“This response misses the point,” an instrumentalist might say, “because the existence of the practice of evaluating people's beliefs as rational or irrational in light of the epistemic goal makes it the case that we can evaluate beliefs in light of the epistemic goal when people do not want to achieve the goal, *and* when we recognize that the goal is not worth achieving. We can make evaluations of instrumental rationality, then, even when the goal in question is neither desired nor good to achieve independent of people's desires.”

But we have to be careful here. It can make sense to evaluate actions and beliefs in light of goals that are neither desired by the agent whose belief is being evaluated, nor good to achieve independent of her desires. Such evaluations can be instructive, for example, because we might be interested in whether it would be instrumentally rational to perform an action in some other possible case where the goal we have in mind is valuable to achieve. Or such an evaluation might tell us something about whether the agent in question is in general trustworthy. But such evaluations are not evaluations of the instrumental rationality of the action or belief at hand, in the service of the goal in

question, in the actual world. They are merely hypothetical evaluations, and do not bear on the actual instrumental rationality of the action or belief in question. Since we are interested in the actual epistemic rationality of beliefs (beliefs are, after all, epistemically rational or irrational in the actual world), merely hypothetical instrumental rationality will not be sufficient.

Notice, by the way, that this restriction on instrumental rationality – that the goal in terms of which evaluations of instrumental rationality take place be valuable in some sense – is quite a weak restriction. In particular, it is neutral between competing views about instrumentalism in the practical domain, where a prominent Humean view is that all that is required in order for a goal to be appropriate to employ in making evaluations of the instrumental rationality of someone's actions is that the goal be desired by that person (e.g. Smith 1995, Hubin 2001). A competing view is that desires are not enough to make a goal appropriate for evaluating the instrumental rationality of actions; the desire must be reasonable, or the goal must be good to achieve (e.g. Korsgaard 1997, Quinn 1993). I do not take a stand on this issue; I am content to allow that a goal can be valuable because an agent desires it, or because of some reason independent of whether anyone desires it. The important thing is that the goal in question in an evaluation of instrumental rationality must be valuable in at least one of these senses.

#### **1.1.4. Some further remarks on instrumental rationality**

A further point needs to be clarified before we wrap up this discussion. On the instrumental conception, the epistemic goal is crucial, but it is not justification/rationality

*itself* that has the epistemic goal. Justification is a status of beliefs; statuses do not in general have goals. Nor is it that *beliefs* have goals; in the words of Pascal Engel, “beliefs do not aim at anything by themselves, they do not contain little archers trying to hit the target of truth with their arrows” (2004, p. 77). Philosophical talk about the aim of belief is best taken metaphorically, to be about the aims that *people* either have or ought to have. (Or one might take it to be about the proper functions of our sub-personal belief-forming cognitive systems, as in Velleman (2000, ch.11) – I consider this suggestion in Chapter 5.) The model for the instrumental rationality of belief in the service of an epistemic goal is the instrumental rationality of action in the service of any other goal we might have (cf. Foley 1987, ch.1). An action is instrumentally rational precisely because it promotes the achievement of a goal that one wants to achieve or that one ought to achieve. The instrumental conception of epistemic rationality takes that view of the instrumental rationality of actions as its model.

To sum up what we have so far, then: there are three broad conceptions of epistemic rationality in which I will be primarily interested. The categorical conception makes epistemic rationality independent of the achievement of an epistemic goal; the subjective instrumental conception makes epistemic rationality depend on the content of the epistemic goals of agents; and the objective instrumental conception makes epistemic rationality depend on the content of an epistemic goal, whether or not people care about achieving it. For the two types of instrumental conception, the goals that matter are goals of agents: goals that agents in fact want to achieve, or else goals that agents ought to

achieve. (There is a third type of instrumentalism, however, which I do not find plausible, which is the hypothetical view, according to which epistemic rationality is instrumental rationality in the service of the epistemic goal, whether or not that goal is valuable to achieve in any sense.<sup>13</sup>)

For quick reference, here is a chart setting out the different varieties of instrumental approach to epistemic rationality:

**Varieties of Instrumentalism about Epistemic Rationality**

Subjective	Epistemic rationality is a matter of holding beliefs that achieve (or are likely to achieve) the epistemic goals that agents care about achieving.
Objective	Epistemic rationality is a matter of holding beliefs that achieve (or are likely to achieve) epistemic goals that are valuable to achieve, whether or not agents in fact want to achieve them.
Hypothetical	Epistemic rationality is hypothetical rationality: if the epistemic goal is valuable in some sense, then it is instrumentally rational to take the means to achieve it. (I will not address the hypothetical approach very much in what follows.)

---

<sup>13</sup> I dislike the use of scarequotes in a written text, but I would like to register an urge to put scarequotes around the word “goal” in that sentence. A goal that is not valuable in any sense to achieve, to my mind, is no goal at all.

One final point is worth emphasizing before moving on. On at least one view of what makes a position deontological, the categorical conception is not the only deontological view. Following Alston (1988), we can think of a deontological approach as one that essentially employs the language of permission, prohibition, and requirement. So understood, each of the three conceptions of epistemic rationality at hand can be fleshed out in deontological terms. For example, the objective instrumental conception might be fleshed out with the claim that it is our epistemic *duty* to always believe in such a way as to have a comprehensive body of beliefs with a favourable truth-falsity ratio. The subjective instrumental conception might be fleshed out with the claim that it is our epistemic duty to always believe in such a way as to promote the achievement of the epistemic goals that we want to achieve. We should therefore avoid talking about the categorical/instrumental distinction in deontological/non-deontological terms.

#### **1.1.5. Reasons and rationality**

The final bit of groundwork to lay down is regarding the relation between reasons and rationality. I take it that, in order for a belief to be epistemically rational, there must be epistemic reasons that support it (i.e. supporting reasons that bear on its truth). Not every belief that has supporting epistemic reasons will be epistemically rational, though, because in some cases the available reasons will conflict. In such cases, because we are interested in rationality in the strong sense, the balance of reasons will determine whether it is rational to believe, to disbelieve, or to suspend judgment.

What is more, when there are good reasons available for a belief, the belief must be *based* on those reasons (in some sense that needs to be fleshed out), if it is to count as epistemically rational. To illustrate what I mean, consider the following cases.

**Jane**

Jane believes that her investments will do very well over the next year. She has that belief because she read it in her horoscope. She is aware of the claims of some prominent economists and bankers who say that the companies she has invested in will do very well, and their stock prices are expected to soar. However, although she knows that economists and bankers are generally reliable about the expected behaviour of stock prices, Jane makes it a point to ignore the advice of economists and bankers, because of the recent worldwide financial disaster. She bases her belief only on her horoscope.

**Sally**

Sally believes that her investments will do very well over the next year. She has that belief because some prominent economists and bankers expect the companies she has invested in to do very well, and their stock prices are expected to soar. Sally has read her horoscope, which agrees with the expectations of the economists and bankers. Sally enjoys horoscopes, but she does not place any epistemic trust in them; the only reason she believes that her investments will do well is the fact that economists and bankers expect it.

Clearly, Sally's belief enjoys some epistemic support which Jane's belief lacks. Without seeing some further details filled in, we are not in a position to make an overall judgement about the epistemic status of the beliefs of Jane and Sally, because the evidence brought forward in the short descriptions of the cases might be overridden by stronger contrary evidence. Still, other things being equal, we can say that Sally's belief is epistemically rational, because it is based on a good epistemic reason. Jane's belief, on the other hand (again, other things being equal), is epistemically irrational, because it is based on a bad epistemic reason. Although Jane is aware of a good epistemic reason, that reason does not support her belief (not even counterfactually; had Jane not read her horoscope, she simply would not have the belief that her investments will do well).<sup>14</sup> The difference between Jane and Sally, then, is that although both are aware of a good epistemic reason for the belief that their investments will do well, only one of them bases her belief on that reason.

Now, it might seem that by requiring that beliefs be based on good reasons in order to be epistemically rational, I am prejudging the issue here in favour of an internalist view of epistemic justification. Internalists, as we saw earlier, hold that whatever it is that makes a belief epistemically justified must be internally accessible, in some sense, to the believing subject. If a subject cannot become aware of some potential justifying reason for his belief, just by reflection, then that potential justifying reason does

---

<sup>14</sup> Or, depending on how the details of the case are filled in, perhaps it *is* rational for Jane to ignore the advice of economists and bankers. However, even if that is the case, Jane's belief that her stocks will do very well is still not epistemically rational; what she epistemically ought to do, if she has no reliable source on which to base a prediction about her stocks, is to suspend belief.

not in fact do any justificatory work.<sup>15</sup> Externalists, as one might expect, deny that claim. The most prominent form of externalism about justification is reliabilism – in its simplest form, the view that a belief is epistemically justified if and only if it is produced by a reliable belief-forming process. And reliabilists might think that the way I am setting up the basing relation rules their position out from the start, because when a reliable belief-forming process is not something that a subject is capable of becoming aware of, he can hardly be said to base his beliefs on the process's reliability.

I admit that this account of the basing relation seems to fit most naturally with an internalist account of epistemic justification, but there are three things to say here. First, internalism about justification seems to me to be correct, so I am not very worried about ruling out reliabilism. Second, emphasizing the importance of the basing relation is well-motivated: when a subject has good reasons available, but fails to form a belief for those reasons, and instead forms a belief for bad reasons, there is something epistemically wrong, and the belief fails to be justified. Perhaps it is *justifiable* for the subject, since he has good reasons available to him, on which he could base his belief, but it is not *justified* unless it is so based. And third, there is a clear sense in which reliabilists (process reliabilists, at least) do respect the basing relation. Because the reliabilist holds that beliefs are justified when they are produced by reliable belief-forming processes, there is a direct connection between beliefs that have reliabilist justification, and the reason that

---

<sup>15</sup> There are other ways to understand what it means for a reason or justifier to be internal to a subject. We can set them aside for now, though, as I am not trying to advance the debate between internalists and externalists here. I am only trying to respond to a potential objection from the externalist camp.

explains why the belief is held. So a form of the basing relation is built right into a reliabilist account of justification.

Let us consider two more cases:

### **Sam**

Sam used to be a chicken sexer. He is able to reliably distinguish male from female chicks. He does not know how he does it; he just does it. Now, Sam was recently fired from his job, and the reason he was given for being fired is that his track record has gone down to the point where he is an entirely unreliable chicken-sexer. In fact, unbeknownst to anyone, Sam is still a reliable chicken-sexer; the data indicating otherwise is the result of a series of clerical errors. However, Sam now believes that he is no longer able to reliably discriminate male from female chicks. Nevertheless, when he goes to visit his brother's farm, and picks up a chick, the belief-forming process that he used to employ as a chicken-sexer generates in him the true belief that the present chick is female.

### **Jim**

Like Sam, Jim used to be a chicken sexer. Also like Sam, Jim was recently fired for being unreliable at distinguishing male from female chicks. And Jim is also in fact still a reliable chicken-sexer; the data indicating otherwise is again the result of a series of clerical errors. Unlike Sam, though, Jim takes chicken-sexing extremely seriously, and as a result, he attempts to learn a new method for distinguishing male from female chicks. (Say, he examines the shape of their

beaks.) He only has flimsy reasons to believe that the new method is reliable, but he believes that it is reliable anyway. And in fact, it is extremely unreliable. Now, when Jim goes to visit his brother's farm, and picks up a chick, his old belief-forming process generates an inclination to believe that the chick is female. Unwilling to trust that inclination to believe, though, he suspends judgment until he examines the shape of the chick's beak. As luck would have it, once he has examined the chick's beak, Jim forms the true belief that the chick is female. So, although Jim has a reliable belief-forming process at hand that could have generated the true belief that he now has, his belief was in fact generated by an unreliable process. Moreover, had Jim not employed the unreliable process, he would not have formed the belief at all, since he did not trust what was in fact the reliable process.

These cases are admittedly a bit complicated. They are set up this way in order to make two things explicit. First, neither Sam nor Jim has internalist justification for their beliefs about the sex of the chick. Sam's belief arises out of a belief-forming process that he has a good reason to think is unreliable – namely, he was fired for being unreliable. Jim's belief arises out of a belief-forming process that, by stipulation, he lacks any good reason to think is reliable (the examination of the chick's beak). Because there is no internalist justification in either of these cases, it follows that if there is some appreciable difference between Sam's and Jim's epistemic situations, that difference must come from an external feature of the cases.

The second thing that ought to come out of the cases is that a reliabilist can say that there *is* an appreciable difference between Sam's and Jim's epistemic positions: although Sam and Jim both have reliable belief-forming processes at their disposal, only Sam's belief is generated by that process. Jim's belief is generated by a different, unreliable process. A reliabilist would therefore say that Sam's belief is epistemically justified,<sup>16</sup> but deny that Jim's belief is similarly justified. And the difference between the two cases looks very much like the difference between a belief that is based on a good reason, and a belief that is based on a poor reason even though there is a good reason ready to hand.

So, even though externalists and internalists disagree about what kind of thing can be an epistemic reason, epistemologists from both camps can accommodate the intuition that, whatever an epistemic reason is, it must be more than merely *available* to an agent in order for it to justify a belief. It must also be the reason *for which* the belief is held.

There is in fact a lot more to say about the epistemic basing relation. Problems arise, among other things, because of deviant causal chains. If we want to say that what it is for a belief *p* to be based on a reason *q* is for *q* to enter into the causal history of the belief *p*, for example, then we will run into cases like the following: you are driving with a hot coffee in hand; you come to believe you are about to be in an accident, which causes you to spill the coffee, burning your hand; and now you have the belief that your hand has been burned. The belief that you would get into an accident certainly figures importantly

---

<sup>16</sup> Or at least that it has some degree of epistemic support – a reliabilist might want to say that its justification is defeated by the fact that Sam has reason to think that the process that generated the belief is unreliable.

in the causal history of the belief that your hand has been burned, but it clearly is not the reason *for which* you believe that your hand has been burned.<sup>17</sup> However, for our purposes here, nothing big hangs on the details of how the account of the basing relation is worked out; what is important is just to see that there is an important difference between holding a belief for a good reason, and having a good reason available but holding a belief for a bad reason instead. Epistemically rational beliefs are those that are held for good reason.<sup>18</sup>

That will suffice by way of laying down the necessary conceptual foundations for our inquiry. What follows now is a preview of what is to come in the following chapters.

## **1.2. A Preview**

In the second chapter, we will briefly go through the views of a number of theorists, both critics and defenders of the instrumental conception of epistemic rationality. There are some serious critics of the instrumental conception, and the view has a number of committed defenders as well. The purpose of this chapter will be to get a feel for what people have to say about the instrumental approach, as well as to establish the importance of inquiring into the overall plausibility of the instrumental conception. There is tension in the literature among epistemologists, centering on the question of the viability of the

---

<sup>17</sup> See Korszcz (1997; 2000) for more good discussion of the basing relation, and some problems about deviant causation.

<sup>18</sup> Sometimes this distinction is marked as the distinction between propositional and doxastic justification. Propositional justification is a matter of there being good reasons for believing a proposition, whether or not the belief is held for those reasons, or whether it is even held at all; doxastic justification is a matter of holding a belief for good reasons. In these terms, I take doxastic justification to be what is most interesting epistemically. We will consider some arguments about the importance of the basing relation in Chapter 6, in the context of an examination of Foley's sufficient-reflection conception of epistemic rationality.

instrumental conception, and that tension needs to be worked out by a thorough examination of instrumentalism in epistemology.

The third chapter is concerned with the reasons that one might offer for accepting the instrumental approach. I discuss five possible reasons for accepting instrumentalism, and I argue that those reasons are not as compelling as they might at first appear. Chief among the reasons for the instrumental approach is naturalism in epistemology. Briefly, the idea is that just about everyone wants to be an epistemological naturalist, and naturalism is committed to some appropriately scientific worldview which does not acknowledge the existence of “spooky” kinds of objects and properties (cf. Mackie 1977). In such a worldview, it is not easy to find a home for normativity. (The basic question: how do you get *oughts* out of *atoms*?) However, desires are parts of the natural world, and ways to achieve desires are easy to talk about naturalistically. The idea is, then, that we can keep a place for talk about epistemic reasons, normativity, and rationality in a naturalist worldview, by talking about such things in instrumentalist terms.

The main reply to this line of argument is that it only appears to be legitimate because it trades on an ambiguity in the meaning of “naturalism.” Metaphysical naturalism, the view that there are no non-physical or otherwise spooky objects and causes, is the kind of naturalism that might be thought to require instrumentalism if we are to retain any talk of normativity and rationality at all. But in epistemology, naturalism comes in all sorts of varieties, not all of which are committed to instrumentalism. Naturalism in epistemology is not a monolithic position that entails instrumentalism.

The point of the third chapter is not to establish that the instrumental conception has nothing going for it. Indeed, it does have some appealing features. But the reasons in favour of it are not as strong as they can appear at first. The conclusion of the chapter is that a healthy step back and re-evaluation of the instrumental conception is in order.

The fourth chapter addresses the formulation of the epistemic goal(s), in order to give the instrumental conception the best hearing that we can. A poor formulation of the epistemic goal(s) will give rise to obvious objections that could otherwise be avoided. This chapter takes up the question of what is of epistemic value, and which epistemic values get to count in the determination of the status of beliefs as epistemically rational or irrational. Obviously, not all things that are epistemically valuable get to count as the goals by reference to which the instrumental conception counts beliefs as epistemically rational or irrational. Having epistemically rational beliefs, for example, is certainly valuable from an epistemic point of view, but that epistemic value cannot count as a goal by reference to which beliefs will count as epistemically rational or irrational, on pain of circularity. It is not so straightforward, then, to identify those values that (for the instrumentalist) will determine the epistemic rationality of beliefs. The formulation of the epistemic goal that I eventually settle on is Foley's: now to believe the truth and now to avoid error.

In the fifth chapter, we get on to the arguments against the instrumental approach. This chapter deals with one way of being an objective instrumentalist: the proper function approach to epistemic rationality (or warrant, entitlement, etc.). A number of

epistemologists approach the analysis of their preferred epistemic concepts by way of the proper function of our cognitive systems. The idea is that what makes it important to achieve the epistemic goal is that it is grounded in the proper function of our cognitive system. This type of view is objectivist, because it holds that the normativity of the epistemic goal is independent of what agents care about or want to achieve. I push a number of objections to the view, most of them centering around two problems: (1) proper functions are not themselves normative; they do not give anyone reasons to do anything, except in special cases where there is some independent reason why fulfilling a proper function is good. And (2) the concept of a proper function makes use of the natural-selection-history of an organ, trait, biological system, etc., but our cognitive systems need not in fact have the right kind of causal history to ensure that they have truth-centered proper functions, or indeed have proper functions of any kind. If they do not, then the proper function approach to epistemic rationality entails that we must be open to the claim that no beliefs are epistemically rational or irrational. But we should not be open to that, at least not without excellent reason – and the fact that a not-very-widely-accepted analysis of epistemic rationality (the proper function approach) entails it, it seems to me, is *not* an excellent reason.

The following three chapters each contain a distinct line of argument against the instrumental conception. The sixth chapter contains an argument to the effect that instrumental rationality in the service of the epistemic goal is not sufficient for epistemic rationality: there are cases of epistemically irrational beliefs that nevertheless are held in

such a way as to promote the achievement of the epistemic goal. The seventh chapter contains an argument based on the value of the epistemic goal: there are cases where there is no sense in which achieving the epistemic goal is valuable, but where there can nevertheless be epistemically rational or irrational beliefs. Such cases are counterexamples to the necessity of instrumental rationality for epistemic rationality. The eighth chapter presents the argument that taking epistemic rationality to be thoroughly instrumental gets us into a vicious regress.

The overall import of the arguments to come is that, although the instrumental approach has some appealing features, those features are not enough to warrant continued adherence to it in light of the serious problems which it faces.

**PART I: Exploring the Instrumental Conception**

## **Chapter 2: Instrumentalists and Their Critics**

The overall purpose of this dissertation is to show that the instrumental conception of epistemic rationality is not tenable. The task of this chapter is to review some literature, in order to establish that the instrumental conception is worth taking seriously enough to merit sustained criticism. Instrumentalism is in fact a fairly widespread view, which appears to have quite a lot going for it. We'll get into the features of instrumentalism that appear to make it an interesting view in Chapter 3, and try to determine how well those features motivate its acceptance. In the present chapter, we will see that there are a number of critics of instrumentalism, in section 2.1, as well as a good number of committed instrumentalists, in section 2.2. In section 2.3, we will briefly look at the views of Michael Lynch and Alvin Goldman, and then the chapter concludes with a final section in which I point to the epistemologists I take to best represent the subjective and the objective instrumental conceptions of epistemic rationality.

### **2.1. Critics of Instrumentalism**

#### **Reynolds**

Steven Reynolds gives a clear, succinct statement of the instrumentalist view, as a possible alternative to his own account of epistemic justification:

A more or less standard account of the relation of truth and justification sees truth as a goal, and justification as an evaluation relative to that goal. One is justified if and only if one believes as one ought. This 'ought' is understood on the model of the 'ought' that occurs in 'If you want go to the market, you ought to turn right at the second stop sign'. It indicates advice about how to achieve a goal. Roughly speaking, the epistemic goal is to acquire true beliefs and avoid false beliefs. (1991, p. 288)

Reynolds does not give much in the way of argument either for or against instrumentalism, instead offering his own account as more plausible than the instrumental conception, but he does take the view seriously enough to consider it as an available alternative to his own view, and his statement of the view is helpfully accurate and succinct.

### **Siegel**

A more serious critic of instrumentalism is Harvey Siegel. Siegel (1989; 1990; 1996a; 1996b) argues against the instrumental conception generally, and against the views of Larry Laudan and Ronald Giere in particular, which he takes to be instrumentalist (and, as we'll see, Laudan and Giere are happy to agree that their views are instrumentalist). Siegel argues, contrary to their views, that “instrumental rationality itself depends on a non-instrumental conception of rationality” (1996a, p. S118), a categorical type of rationality; he holds that instrumental rationality of any kind always depends on an underlying categorical epistemic rationality. Briefly, the idea is that in order for it to be instrumentally rational to adopt any means to achieve a goal, the available evidence must make it epistemically rational to think that the means are good ones for achieving the goal. I develop this line of argument in Chapter 8.

### **Fumerton**

In his (2001a), Richard Fumerton spends a few pages arguing that the instrumental construal of epistemic judgments is mistaken. He explains the goal-oriented conception of epistemic judgments:

On one (rather crude) view, what one prudentially ought to do is what maximizes satisfaction of one's desires. What one legally ought to do or what one is legally justified in doing is a function of the extent to which an action satisfies the goal of following the law. What one ought to do from the standpoint of etiquette is a function of following the goals or ends set down by the 'experts' who worry about such things. So all one has to do to fit the epistemic "ought" into this framework... is delineate the relevant goals or ends that define what one epistemically ought to believe. And the obvious candidates are the dual goals of believing what is true and avoiding believing what is false...

Now, as plausible and potentially illuminating as this account might seem initially, it is, I think, fatally flawed. (2001a, pp. 54-55)

Fumerton finds the instrumental conception interesting, but he thinks that it does not succeed. I will not dwell on Fumerton's arguments against instrumentalism, though, since other critics of instrumentalism develop more or less the same lines of criticism at greater length and in greater detail.

### **Kelly**

A serious critic of instrumentalism is Thomas Kelly. Kelly's explicit target is what I call the subjective instrumental conception of epistemic rationality: the view that "epistemic rationality is a species of instrumental rationality, viz. instrumental rationality in the service of one's *cognitive* or *epistemic* goals" (2003, p. 612, emphasis in original). The kind of epistemic goals that Kelly has in mind are the goals that agents in fact care about achieving, so his main target is the subjective instrumental conception. He has a two-pronged argument against instrumentalism, one prong of which I will make use of in Chapter 7. Briefly, the argument is that there are cases where one does not want to achieve the epistemic goal, and in such cases, it cannot be instrumentally rational to take the means to achieve it (i.e. to acquire true beliefs and avoid false ones). Kelly also

recognizes that one might opt for the objective instrumental conception instead, and he gives a brief argument against that view as well (2003, pp. 632-33). We will come back to Kelly's arguments later.

## **Grimm**

Finally, Stephen Grimm considers the instrumental conception to be a prominent one, and he has recently argued against it. He writes:

Among contemporary epistemologists, perhaps the most prominent way to make sense of our epistemic evaluations is in teleological terms. On this way of looking at things, a belief earns positive marks, from an epistemic point of view, just to the extent that it seems to promote or in some way bring about the things with intrinsic epistemic value. And similarly, a belief earns negative marks just to the extent that it seems to *fail* to promote or bring about the things with intrinsic epistemic value. ...one of my basic goals in this paper will be to show that the teleological view – at least, as it is popularly understood – is mistaken. (2009, p. 243; see also his 2008.)

The fact that these theorists, who are not themselves instrumentalists, consider the instrumentalist view an important one, and worthy of consideration and refutation, indicates that the view is an important one, and worthy of being taken seriously.

## **2.2. Defenders of Instrumentalism**

In this section, we'll see that there are a number of epistemologists who defend instrumentalism of one sort or another. Some explicitly espouse the subjective instrumental conception, while others take up the objective conception. Some others are not explicit about which type of instrumentalism they prefer. What follows is a list of epistemologists who have defended instrumentalism in one form or another.

## Leite

Although he is not himself either a serious critic of instrumentalism or a committed instrumentalist, Adam Leite responds to Kelly's argument on behalf of the instrumentalist. He writes:

According to one popular account, epistemic rationality is a species of instrumental rationality: a belief is epistemically rational when (and because) holding it is instrumentally rational given one's cognitive or epistemic goals; and one has an epistemic reason to believe something when (and because) doing so would be instrumentally rational given those same goals...

I am not myself a committed instrumentalist. However, I think that there is an important response open to the instrumentalist, and I would like to develop it in some detail. (2007, p. 456)

## Laudan

Larry Laudan explicitly defends an instrumental conception of reasons in general, as well as of epistemic reasons in particular. He writes, for example, that "Good reasons are instrumental reasons; there are no other sort" (1990, p. 320). He also writes, in setting up his response to Harvey Siegel's objections, that

Siegel has many criticisms of my approach but most of them boil down to Siegel's insistence that instrumental rationality – of which he rightly takes me to be an advocate – and epistemic rationality are fundamentally different things... As I shall try to show here, Siegel (1) fundamentally underestimates the resources of instrumental rationality, (2) fails to see that epistemic rationality is a species of the genus instrumental rationality... (1990, p. 316)

So Laudan is quite explicit about being happy to think of himself as an instrumentalist about epistemic rationality. Further on, he writes that

Justification is itself a *relational* notion. To say that 'x is justified in doing y' is always enthymematic for 'x is justified relative to end(s) in doing y'. There is no coherent sense of justification (epistemic or otherwise), just as there is no sense of deliberative action (epistemic or otherwise), in the absence of the specification of

the ends with respect to which an action is deemed justified or rational. That is the central premise of instrumental rationality and of normative naturalism. (1990, p. 317)<sup>19</sup>

Laudan endorses the instrumental conception of epistemic rationality here, because he holds that no one can have an epistemic reason in the absence of an epistemic goal. His main reason for endorsing the instrumental view of epistemic rationality is that he thinks that there simply is no other coherent view of rationality (epistemic or otherwise) to be had.

Finally, Laudan's instrumentalism appears to be subjectivist, in the sense that the goals that are important for epistemic purposes are *desired* goals: "The theory of instrumental rationality simply insists that, once one has settled on one's cognitive utilities or desired ends, then the issue of the appropriate methods of appraisal to use depends on what strategies conduce to the realization of the selected end" (1990, p. 318).

To be clear, the point is not that Laudan denies that we can make epistemic judgments without adopting an epistemic goal. We can make means-ends judgments about ends that we do not care about, epistemic ones included, and Laudan has no problem with that. The point is, rather, that a subject cannot *have* an epistemic reason to adopt a belief, nor can her beliefs be epistemically rational, on this conception, unless the subject has an epistemic goal.

---

<sup>19</sup> Notice that Laudan is using the notions of justification and rationality interchangeably here.

## Giere

Like Laudan, Ronald Giere argues for a naturalized philosophy of science. Distinguishing instrumentalism from the categorical conception of rationality, he writes:

[T]here is another, weaker, form of rationality which is conditional, or *instrumental*. To be instrumentally rational is simply to employ means believed to be conducive to achieving desired goals... Finally, there is also a more *objective* sense of instrumental rationality which consists in employing means that are not only believed to be, but are *in fact* conducive to achieving desired goals.

This latter, objective, sense of instrumental rationality provides the naturalist theorist of science with ample means for making normative claims about science. (1989, p. 382)

And further on:

Thus a naturalized philosophy of science... can provide a basis for *normative* judgments. These judgments would, of course, be only *instrumentally*, and not *categorically* normative. But for the naturalist, that is the only kind of normative judgment anyone can make. There is no “higher” rationality. (ibid.)

So for Giere a naturalist can only appeal to instrumental considerations. Unlike Laudan, he does not claim that the categorical conception of rationality (epistemic or otherwise) is incoherent, but he does believe that it is incorrect. Giere is a naturalist, and he holds that naturalism can only appeal to considerations of instrumental rationality.

It might be objected, on Giere’s behalf, that what he is interested in is a concept of *scientific* rationality, rather than epistemic rationality. We therefore cannot lump Giere in with the epistemic instrumentalists who are the target of this dissertation. However, there are three points to make regarding that objection. First of all, it is not clear what the distinction is supposed to be between scientific and epistemic rationality. Scientific rationality is widely held to be the paradigm of epistemic rationality: scientists (ideally)

follow careful methods, they check their results, they only draw conclusions based on sufficient evidence, etc. It is not clear that scientific rationality and epistemic rationality come apart, except insofar as the standards of science are higher and more rigorous than the everyday standards of epistemic rationality which apply to all of us. That higher degree of rigour does not make scientific rationality into a different *kind* of rationality than the epistemic sort; it is simply a rigorous species thereof.

The second thing to say to the objection is that Giere does not appear to want to distinguish epistemic from scientific norms. In a section entitled “Naturalistic Justification of Epistemological Norms,” for example, he defends a descriptive approach to epistemology, and he talks about the “epistemological norms of science” (2001, p. 58). And, in that same section, he writes that

One must distinguish two kinds of norms, *categorical* and *conditional*. Categorical norms proscribe or prescribe various actions unconditionally. They simply say, “Do this” or “Do not do that.” Conditional norms have the form: “If you want to achieve G, do A.” ... Conditional norms... can be *justified* naturalistically, and science requires only conditional norms. (2001, pp. 57-58; see also his 2006, esp. pp. 54-56)

Because he is quite explicit about taking the epistemic norms that underlie science to be instrumental, or conditional – that is, of the form “if you want to achieve G, do A” – it is fairly clear that Giere wants to adopt the instrumental approach to epistemic rationality.

And finally, if that is not enough, notice that Giere (1989) takes the trouble to respond to Siegel’s criticism of his instrumentalism, not with a complaint that Siegel is misinterpreting him when he claims that Giere is an instrumentalist about the epistemic norms that underlie the scientific enterprise, but rather with arguments to the effect that

the epistemic norms of science really are instrumental. It is therefore fair to take Giere to be defending the instrumental conception of epistemic rationality.

One point should be emphasized here before moving on. Giere is cited above as holding that the “objective” sense of instrumental rationality provides naturalism with the tools to make normative judgments. He is not using “objective” in the sense that I am here. My sense of “objective,” in the phrase “objective instrumental conception of epistemic rationality,” indicates that the goals with respect to which an instrumentalist makes judgments of epistemic rationality are independent of the actual goals that agents value or desire to achieve. Giere, on the other hand, uses the term “objective” to indicate that the means chosen to achieve the goals are those that will *in fact* be effective, as opposed to those that are merely *believed* by the subject in question to be effective, to achieving the goals. Giere still holds that the goals in question are *desired* goals, so his view is subjectivist, in my sense of the term. Throughout this dissertation, when I write “objectivism” without qualifying it, I mean objectivism about the epistemic goal. When I need to make a point about objectivism and subjectivism about the *means* required to achieve the goal, I will make that explicit.

### **Foley**

Richard Foley is certainly one of the most explicit recent epistemologists about his adoption of the instrumental conception of epistemic rationality. Foley’s view, like Giere’s, appears to be subjectivist, in the sense that the goals that matter for assessments of rationality are goals that agents care about achieving. He claims to be open to an

objectivist reading of epistemic goals as well (1993, p. 4), but he develops the account in subjectivist terms. Unlike Giere, though, who is an objectivist in the sense that he is interested in the means that would *in fact* be conducive to achieving desired goals, Foley is a subjectivist in that sense as well: he holds that it is the means that the subject would, on reflection, take to promote the achievement of the goal that count in epistemic evaluations.

For more than two decades now, Foley has been expounding his instrumentalist view. In his (1987), for example, he explains his view of the general form that rationality takes: “the form rationality takes in the simplest cases is as follows: A person has a goal X, on careful reflection he would believe Y to be an effective and nontrivial means to X, and he brings about Y” (ibid. p. 6). Different types of rationality are distinguished, on this view, by the goals that they involve. Accordingly, epistemic rationality is distinguished from other forms of rationality by virtue of the distinctly epistemic goal at issue: “if we are interested in identifying a distinctly epistemic kind of rationality, it is necessary to identify a distinctly epistemic goal” (ibid. p. 7), and the epistemic goal is understood as “now to believe those propositions that are true and now not to believe those propositions that are false” (ibid. p. 8).

Foley sticks with that view of epistemic rationality through his (1993) and (2001). He writes, for example:

my proposal is that it is best to think about judgments of rationality as judgments concerning how effectively individuals or groups are pursuing their goals. This is so regardless of what it is that we are rationally evaluating... A desideratum of a theory of rationality is that it provide a uniform way of thinking about all such

questions. Rational belief, for example, should not turn out to be a fundamentally different phenomenon from rational action, as if the two shared only a common name. A goal-based approach to rationality, as I will try to show, can satisfy this desideratum. (1993, p. 4)

The only salient change that Foley's view undergoes from his (1987) onward, it appears, is that the conception of the epistemic goal is slightly revised.<sup>20</sup> In his (1987), again, the goal was understood as the goal of now believing truths and now not believing falsehoods. By his (1993), he adopts a slightly broader conception of epistemic goals: any goal that is "concerned solely with the accuracy and comprehensiveness of our current belief system" is a purely epistemic goal (1993, p. 19). The goal of now believing truths and now not believing falsehoods is one variation on this goal, and it is the one that Foley continues to work with, but he is open to other variations as well.

More recently, in his (2008), Foley again explains his take on rationality generally, and epistemic rationality specifically. The "general template of rationality" is: "an action A (or decision, plan, intention, strategy, belief, etc.) is rational for a subject S if it is rational for S to believe that A would acceptably satisfy her goals" (ibid. p. 45).

Noting again that there are many different goals that people pursue, Foley refines this general notion of rationality as follows: "an action A (or decision, plan, intention,

---

<sup>20</sup> Most of Foley's view remains unchanged, e.g. the view of epistemically rational belief as what a subject would on careful reflection take to achieve the epistemic goal, and the view of sufficient reflection from the epistemic point of view as reflection up to the point beyond which thinking any more will not change one's mind (1987, p. 35; 1993, p. 99). One difference, though, is that in his (1993), the subjective foundationalist conception of epistemically rational beliefs disappears. Subjective foundationalism, as Foley develops it in his (1987), is the view that epistemically rational beliefs are those beliefs either that one would on careful reflection take to be properly basic, or else one would take oneself on careful reflection to have an argument for the belief, where this argument is itself based on properly basic beliefs, and one would take the argument to be truth-preserving.

strategy, etc.) is rational in sense X for S if it is rational for S to believe that A will do an acceptably good job of satisfying her goals of type X” (ibid. p. 46). Further on (ibid. p. 48), Foley further refines the general notion of rationality, by specifying that it must be *epistemically* rational for S to believe that A will do an acceptably good job of achieving her goals, in order for it to be rational for S to do A.

To get an account of distinctively epistemic rationality, on this picture, we input the epistemic goal – the goal of “now having accurate and comprehensive beliefs” (ibid. p. 50) – into the general template of rationality, which yields the following characterization of epistemic rationality: “Believing P is rational in an epistemic sense if it is epistemically rational for S to believe that believing P would acceptably contribute to the epistemic goal of S’s now having accurate and comprehensive beliefs” (ibid. p. 54). That account of epistemic rationality admittedly looks circular – epistemic rationality is defined in terms of what it is epistemically rational for S to believe – but it looks like Foley can escape the charge of circularity, by way of his account of sufficient reflection. We can postpone a discussion of Foley’s sufficient-reflection view of epistemic rationality until Chapter 6, though, since the point at issue here is only whether Foley is in fact an instrumentalist about epistemic rationality, and it is quite clear from the passages cited that he is.

Before moving on, I want to note that although Foley is a paradigm subjectivist about the epistemic goal, insofar as his account of rationality involves goals that people want to achieve, he is open to the possibility that people might not want to achieve the

epistemic goal. Even if they do not want to do so, however, Foley suggests that we all might be committed in some way to achieving it. So when he is pushed on the question of desiring to achieve the epistemic goal, Foley might want to shift into an objectivist account of epistemic rationality (e.g. 1987, pp.11-12).

### **Alston**

In his earlier work in epistemology, William Alston was an instrumentalist about epistemic justification. He wrote, for example, that “our central cognitive aim is to amass a large body of beliefs with a favourable truth-falsity ratio. For a belief to be epistemically justified is for it, somehow, to be awarded high marks relative to that aim” (1985, p.59). That is to say, on this view, epistemic justification consists of doing well, considered in light of the epistemic goal (“central cognitive aim”) of having a favourable truth-falsity ratio in a large body of beliefs.

Alston has since taken up the view that there is no single concept of epistemic justification that epistemologists are all trying to pin down. In his (2005), he holds that there are a number of different epistemically important concepts, and that they are not truly competing accounts, because epistemologists have different cognitive goals or epistemic desiderata in mind when they develop them. He takes this no-justification view to offer the best explanation of the entrenched, widespread disagreement about some of the most basic features of justification.

However, although Alston is no longer interested in developing an account of epistemic justification, because he thinks that there is no single property that justification

could be, his position remains instrumentalist in nature, and it is threatening to a non-instrumentalist in that the various ways of performing epistemic evaluations according to his pluralist account are still relative to the most basic epistemic goal. He writes:

We evaluate something epistemically... when we judge it to be more or less good or bad from the epistemic point of view, that is, for the attainment of epistemic purposes...

The evaluative aspect of epistemology involves an attempt to identify ways in which the conduct and products of our cognitive activities can be better or worse vis-à-vis the goals of cognition. And what are those goals? Along with many other epistemologists I suggest that the primary function of cognition in human life is to acquire true rather than false beliefs about matters of interest or importance to us. (2005, p.29)

Alston varies the formulation of the epistemic goal, characterizing it variously as “*the acquisition, retention, and use of true beliefs about matters of interest and/or importance*” (ibid. p.30, emphasis original), or “maximizing true beliefs and minimizing false beliefs about matters of interest and importance” (ibid. p.32). He also suggests that we might think of the basic epistemic goal as “with respect to any proposition that is of interest or importance to us, to believe it if and only if it is true” (ibid.). The details vary, but the basic idea remains the same, and it remains the case that, on his account, the various epistemic statuses are all evaluations that are somehow positively or negatively related to the basic epistemic goal.

Alston does not restrict our epistemic goals to the one that he identifies as the most basic goal. He allows that there are other epistemic goals and desiderata, but they are parasitic on the basic truth-directed epistemic goal:

...truth is not the sole desirable feature of belief from the epistemic point of view, defined in terms of the primary aim of cognition at true belief. The crucial point is

that the most *basic* aim of cognition is not the only thing aimed at by cognition, not even the only thing aimed at from the standpoint of that most basic aim. That is because other features of belief are also desirable from the standpoint of that most basic aim. (2005, p.36)

The basic point to take away from this brief summary of Alston's view is that, although in his later work he is no longer interested in an account of epistemic justification or rationality, he is interested in the various types of epistemic evaluation, and on his view, the various types of epistemic evaluation all count as epistemic by way of being related to the epistemic goal. Alston's pluralist view therefore remains incompatible with any non-instrumentalist type of epistemic evaluation.

### **Tannsjö**

Torbjörn Tannsjö adopts an explicitly instrumentalist account of epistemic rationality and norms: "To be justified is to hold beliefs that it is rational to hold, given the rest of one's beliefs together with a desire which can be characterised as a desire to realise a specific epistemic goal" (2010, p.105). Tannsjö does not want to argue that this statement captures the single correct meaning of the term "epistemic justification," but he does think that it is the most fruitful account of epistemic justification, and he likes it because he thinks it is a good way to avoid appealing to any *sui generis* epistemic norms. He roughly characterizes the distinction between epistemic and other reasons as follows: "First of all, a person, S, has epistemic reasons to believe a proposition, p, if his or her belief in p is (subjectively) rational in view of his or her epistemic goal. And, secondly, the epistemic goal of a person is the interest this person has in holding the belief the content of it is true [sic], and in not holding it if it is false" (ibid., p.109). If a person desires to achieve the

epistemic goal, which Tannsjö takes to be wanting to have the belief if it is true and not wanting to hold it if it is false, then she has epistemic reasons to form or refrain from forming the belief in question.

### **Maitzen**

Stephen Maitzen (1995) raises an objection to what he calls the “nominal aim” of epistemic justification. The nominal aim that Maitzen has in mind is the epistemic goal adopted by the likes of Alston (1985) and Foley (1987), the goal (roughly) of having true beliefs and not having false beliefs. Maitzen’s objection is that, given an epistemic goal of this kind, there is no room for justified false beliefs or unjustified true beliefs, since any true belief will necessarily promote it, and any false belief will necessary hinder it.

In light of this objection, Maitzen does not propose to abandon the instrumental conception, though. He sees no problem with instrumentalism about justification. Rather, what he proposes is that the nominal aim must be mistaken; the epistemic goal needs to be formulated differently (1995, p.875).

### **Vahid**

Hamid Vahid notes that it is common to take epistemic rationality to have to do with achieving an epistemic goal:

It is generally thought that there is an intimate connection between justification and truth. This is usually construed along the lines that a belief is justified if forming that belief is a good thing from the epistemic point of view which, in turn, is characterized in terms of the aim of maximizing truth and minimizing falsehood in a large body of beliefs. (2003, pp.1-2)

He proceeds to defend the truth-directed nature of the epistemic goal from objections that have been raised against it. In particular, he defends it against Maitzen's objection (which, again, is that the "nominal aim" of achieving true beliefs and avoiding false ones entails that all true beliefs are justified and all false beliefs are unjustified, which is an unacceptable result). Vahid's defense consists of a reformulation of the truth-goal, so that the goal is diachronic rather than synchronic: "the truth-directed goal is a historical concept. It is the overall maximization of truth and minimization of falsity in one's belief repertoire in the long run that it is concerned with" (ibid. p.86). This diachronic formulation of the epistemic goal allows Vahid to hold that beliefs are only appropriately related to the epistemic goal, and therefore only justified, if they have truth-conducive causal histories (i.e. if they are produced by reliable processes, or if they are based on adequate grounds). Therefore, for Vahid, there is room for justified false beliefs and unjustified true beliefs, on a truth-directed instrumentalist conception, since false beliefs can have truth-conducive causal histories and true beliefs can have causal histories that are not truth-conducive.

It need not concern us for the moment whether Vahid's move of making the epistemic goal diachronic is required in order to avoid Maitzen's objection, or whether taking the epistemic goal to be diachronic is independently plausible. What is important for now is that Vahid accepts the instrumental conception of justification – quite explicitly so: "beliefs are rational (justified) to the extent that they serve the epistemic goal of believing truth and not believing falsehood" (2006, p.320) – and he defends a

particular formulation of the epistemic goal, in order to allow the instrumental conception to avoid getting the extension of epistemically justified and unjustified beliefs wrong.

## David

Marian David follows Alston's early take on the epistemic goal, loosely characterizing it as "the goal of *believing truths and not believing falsehoods* (2001, p.152, emphasis in original), or as "the goal of believing *p* if and only if *p* is true" (ibid. p.153). Noting that the standard account of knowledge involves belief, truth, and (Gettier-defeating) justification, David claims that the main target of epistemological theorizing is epistemic justification, because truth and belief are non-epistemic concepts. He also notes that epistemologists want to try to give an account of epistemic justification in non-epistemic terms, to "provide a nonepistemic 'anchor' for justification by connecting it in some significant manner with nonepistemic concepts" (ibid. p.154). The truth-goal, on this picture, provides that anchor. David continues:

Why is truth typically cast as a *goal* when this connection is made? Alston provides the reason. It is generally agreed that being justified is an evaluative concept of some sort: To say that believing *p* is justified or unjustified is to evaluate believing *p*, in some sense, as being a good thing or a bad thing, as having some positive or negative status. The suggestion is that this type of evaluation, epistemic evaluation, is most naturally understood along broadly teleological lines, as evaluating beliefs relative to the standard, or goal, of believing truth and avoiding error. (2001, p.154, emphasis in original)

David also provides a helpful clarification regarding the invocation of the truth-goal in a teleological account of epistemic justification, making it explicit that (as we saw in Chapter 1) the concern with a truth-goal is not merely that we *care* about or *value* justification because it gets us truths. Rather, the truth-goal is part of what *constitutes*

justification. He writes: “The truth-goal can be invoked at two different levels: first, to characterize the nature of justification in broadly teleological terms... second, to explain why we *value* justification, why we care whether our beliefs are justified” (2001, p.161). David argues that Maitzen (1995) slides between these two ways of invoking the truth-goal in a problematic fashion. The first way of invoking the truth-goal is the only one that is at issue in giving a teleological account of the nature of epistemic justification.<sup>21</sup>

Before moving on, I want to point out that David is not clear on whether he is a subjectivist or objectivist with respect to the epistemic goal, or whether he wants to try to chart a middle path between the two. He suggests a number of different takes on the question. His most sustained line of thought appears to be the following. Roughly, the idea is that the epistemic goal need not be one that any particular agent has, in the sense of desiring to achieve it, as long as there are a sufficient number of epistemic agents who do have the epistemic goal and who engage in the practice of epistemic evaluation relative to that goal. Where (G) stands for the goal-oriented conception of epistemic justification, and (D) stands for the thesis that an agent must desire to achieve true beliefs and avoid falsehoods at time *t* in order for her beliefs to be subject to epistemic evaluation at *t* (2001, p.155), David writes:

Distinguish the (nonnatural) evaluative *concept* of justification from the *natural* (nonevaluative) *property*, *F*, that the concept refers to, or picks out, or supervenes

---

<sup>21</sup> In Maitzen’s defense, a passage from BonJour (1985, pp. 7-8) that is widely cited by both friends and foes of instrumentalism, and to which Maitzen himself appeals in explaining the view at issue, lends itself to blurring the line between those two ways of invoking the truth-goal. I will refrain from characterizing BonJour as an instrumentalist precisely because the passage in question is ambiguous between those two ways of invoking the truth-goal. Furthermore, BonJour’s (2010) is more explicit about taking truth to be a goal in the second sense: it is why we *value* justification, but not part of what constitutes justification.

on. The concept of justification as an evaluative concept exists only because we, or enough of us, engage in the practice of evaluating beliefs in epistemic terms. So (G) is indeed committed to (D), with the consequence that our beliefs wouldn't be justified, that is, wouldn't fall under the concept of justification, if no one desired having true beliefs, because there would be no such evaluative concept for beliefs to fall under. (2001, p.157)

However, this take on the subjective/objective distinction is in fact a red herring – it distracts attention from the real point at issue, which is whether the goals that matter for epistemic evaluation must be those that *the agent in question* desires to achieve, or not.<sup>22</sup> Or another way to put the point is: the content of (D) has to change from its initial individualistic formulation (2001, p.155) to a group-consensus formulation of some sort, if this passage is to make any sense, so that this proposal is irrelevant to the initial question. (David does not explicitly say that the view proposed in the passage just cited is his view regarding the relation of theses (G) and (D), but it appears to be his favoured view.<sup>23</sup>)

Finally, I want to offer a brief digression on David's eventual formulation of the epistemic goal, which is given in subjunctive terms: "For every  $p$ , if I were to believe  $p$ , then  $p$  would be true, and if  $p$  were true, then I would believe  $p$ " (2001, p.166). This is the goal that David takes to be promoted by reliabilism. He makes this point in response to

---

<sup>22</sup> Perhaps, instead of merely offering a red herring, David is defending a version of the alternate, *hypothetical* type of instrumental evaluation which we saw in Chapter 1 – the view that evaluations in light of a goal can make sense even if the goal is not valuable itself, nor is it desired by the agent in question, as long as enough people care about the goal to get the practice of evaluating beliefs or actions in light of that goal off the ground. But (1) if that is his point, then it distorts the initial statement of the thesis (D), which was that a subject must herself desire to achieve the epistemic goal. And (2), if this is David's point, then it will be subject to the objections raised in Chapter 1.

<sup>23</sup> However, he also writes: "the goal-oriented approach to justification may get by without the thesis that we actually desire truth" (2001, p. 160). And he points out (citing Foley 1993, p. 17) that perhaps there is a counterfactual sense in which agents *would* on reflection adopt the truth-goal, so he is not necessarily committing himself to any particular position regarding theses (G) and (D).

Maitzen's (1995) objection to reliabilism, that it abandons the truth-goal in favour of the goal of having reliably produced beliefs. David's point is that this subjunctive formulation of the truth-goal *is* promoted by reliabilism, because to say that a belief is reliably produced is to say that it belongs to a family of beliefs, most of which are such that, if they are believed, then they are true.

Note, though, that this subjunctive goal is not necessarily promoted by Goldman's classic *process* reliabilism. It is perhaps promoted by some sort of *belief* reliabilism, such as Nozick (1981)-style tracking accounts (the combination of theses: if S believes *p*, then *p* is true; if *p* is false, then S doesn't believe *p*), or Sosa's (2007) basis-relative safety (if S believes *p* on *this* kind of basis, then *p* is true in most close possible worlds in which the belief that *p* is so based) – but even that much is not certain. Goldman (1986, ch.5) takes great pains to argue that logic (among other things) cannot by itself generate right justification-rules, because people can make valid inferences in unreliable ways, i.e. by instantiating unreliable psychological processes. For Goldman, it is the process that counts. But notice, now, that S can satisfy the goal “if *p* were true, then S would believe *p*” by the use of *unreliable* processes.<sup>24</sup> Suppose that S has a number of belief-forming processes available, most of which are unreliable, but all of which would generate the belief that *p* in any close possible world involving the case at hand. Suppose also that *p* is true, and that process P1, an unreliable process, generates the belief that *p* in

---

<sup>24</sup> This is not a knock against Goldman, since he explicitly is not interested in how well individual beliefs fare with respect to the epistemic goal. He is a rule-consequentialist; it is the rules that are primarily justified by reference to the goal, and beliefs are indirectly justified by way of being produced by such rules. This is only a knock against David's defense of reliabilism's claim to being committed to the truth-goal, against Maitzen's objection.

S. Then it follows that S achieves the subjunctive epistemic goal ( $p$  is true, and S believes  $p$ ), without having produced the belief by a reliable process. So there can be beliefs that achieve the subjunctive epistemic goal, that are not reliably produced. Conversely, one can also have reliably-produced beliefs that do not achieve the subjunctive epistemic goal. Suppose that process P2, a reliable process, generates the false belief that  $q$  in S. Then it follows that S has a reliably-produced belief, but it does not achieve the subjunctive truth-goal, i.e., it is not a belief that S would have only if it were true.

All of that is to say: the subjunctive truth-goal – that S would believe  $p$  if  $p$  were true, and that S would not believe  $p$  if  $p$  were false – is not necessarily a goal that is promoted by reliable belief-forming processes.

### **White**

We find another epistemologist making use of the instrumental conception of epistemic rationality in Roger White. He writes: “The three themes in contemporary epistemology that I will be drawing on are *instrumentalism*, *internalism*, and *conservatism*” (2007, p.117). In characterizing instrumentalism, he writes:

It is natural to suppose that *epistemic* evaluation is an evaluation of how we are doing with respect to our goal of finding the truth. So the thought is that epistemic rationality is to be understood along the lines of instrumental or means-ends rationality. The rough idea is that the epistemic rationality of a belief is a matter of how the means of belief formation is conducive to the end of obtaining the truth. (2007, p.117)

White proceeds to attempt to resolve some apparent problems that face instrumentalism. Of importance for us is that he gives a hypotheticalist response to Thomas Kelly's arguments against instrumentalism. We will return to that response in Chapter 7.

### **Riggs**

Wayne Riggs's concern is Jamesian in spirit. He is concerned with the balance between the goal of achieving true beliefs and the goal of avoiding falsehoods. Riggs argues that the fact that these two goals really are not equivalent requires that epistemologists give a great deal more attention to the formulation of the epistemic goal than they have done; we cannot go on treating the goal as a single truth-goal, as though its constituents cannot come apart.

Epistemic justification is conceptually tied to our epistemic goals, for Riggs: "no matter how one defines 'epistemic justification,' it won't be *epistemic* justification unless there is some explicit conceptual tie between the evaluation being defined and the goals that we take to be 'epistemic' ones" (2003, p.345, author's emphasis). Furthermore, Riggs argues that the weight that we assign to the two goals will determine the epistemic rationality of our beliefs. Riggs is working with just about the most explicitly instrumental conception of epistemic rationality that there could be, an expected-utility account of epistemic rationality: "the epistemic rationality of a given belief is assessed in terms of its expected utility with respect to the goals of having true beliefs and avoiding error" (2008, p.6).

Riggs's conception combines elements of subjectivism and objectivism. He holds that it is the way that subjects weigh the epistemic goals of achieving true beliefs and avoiding false beliefs that makes their beliefs epistemically rational or irrational. But he also holds that there are limits on rational ways of weighing the epistemic goals (one may not value avoiding falsehoods exclusively and still remain epistemically rational, for example), and he also holds that achieving truths and avoiding falsehoods are the only candidate epistemic goals.

### **Steglich-Petersen**

Finally, I want to point out that Asbjørn Steglich-Petersen has an interesting recent defense of an instrumental account of epistemic reasons. The basic notion of an instrumental or teleological account, for Steglich-Petersen, is as follows: “on the teleological conception of reasons for belief, whether someone has reason to believe some particular proposition *p* on some particular occasion, depends on the value of the result of believing *p*, or the intrinsic value of believing *p*, on that occasion” (2011, p.13). One important challenge that instrumental accounts face is that there does not seem to be any universally valuable epistemic property. (We will consider some points that support this challenge in Chapter 7.) Steglich-Petersen concedes the challenge – a teleological account of epistemic reasons that relies on the claim that there is an epistemic property that is valuable in all contexts is bound to fail – but he does not think that that is the end of instrumentalism. He continues, “So the question becomes if a teleological explanation of epistemic reasons, which doesn't rely on that claim, is available. I will introduce and

motivate such an account” (ibid. p.15). We will discuss Steglich-Petersen’s proposal in Chapter 7.

### **2.3. Addendum**

Finally, I want to draw attention to the views of Michael Lynch and Alvin Goldman, even though they are not instrumentalists of the sort that I am primarily concerned to respond to in this dissertation. Then I will close the chapter by identifying the principal instrumentalists whose views I take, throughout much of the rest of the dissertation, to best represent the subjective and the objective instrumental positions.

#### **Lynch**

Michael Lynch is not interested in the instrumental conception of epistemic rationality, as far as I can tell. His concern is not with whether the epistemic goal has anything to do with what makes beliefs epistemically rational or irrational. However, he is an explicit advocate of the objective value of the goal of having true beliefs and avoiding false beliefs: “When I say that true belief is a “proper” end of inquiry, I mean that true belief is something that is worth pursuing, whether we in fact desire to pursue it” (2009, p.77). So, although Lynch is not an explicit advocate of the instrumental conception of epistemic rationality, he is an objectivist about the value of truth, and if objectivism about the value of truth is correct, that goes a long way toward establishing the plausibility of the instrumentalist conception.

## **Goldman**

In his “individual” epistemology (i.e. his account of belief-justification), Alvin Goldman holds that truth is intrinsically valuable: “Truth acquisition is often desired and enjoyed for its own sake, not for ulterior ends. It would hardly be surprising, then, that intellectual norms should incorporate true belief as an autonomous value, quite apart from its contribution to biological or practical ends” (1986, p.98). In his later social epistemology (his account of the epistemic value of various social institutions and practices), he holds that it is only *interesting* true beliefs that are of intellectual value – either beliefs that are interesting to the subject in question, or beliefs that a subject has that others are interested in (1999, pp.88-89). For the purpose of an account of justification, I will treat his take on truth in his individual epistemology to be his settled view.

The structure of Goldman’s account of justification involves a system of justification-rules, which either permit or require the formation of doxastic attitudes. The rules themselves are justified by reference to the truth-goal. Beliefs are justified derivatively, by way of being produced in accord with right justification rules; the consequences of individual beliefs with respect to the truth-goal (or in any other respect) do not bear on their epistemic status: “Of course I restrict myself to *rule* consequentialisms, for I am interested in the rightness of rules (or rule systems). I ignore entirely the suggestion that the justificational status of each belief is a function of that very belief’s consequences” (1986, p.97, emphasis in original). Although Goldman’s view differs from many of the instrumentalists we have been considering, because he

holds that whatever the consequences of a belief may be with respect to the epistemic goal, those consequences do not enter into the determination of that belief's epistemic status, it is still importantly instrumentalist, because justification depends on an indirect relation to the epistemic goal. Now, the main target of this dissertation is not the rule-consequentialist approach of people like Goldman, but we will see two objections that will apply to Goldman's view, in Chapters 6 and 7. Briefly, the first is that there are cases where subjects can form beliefs in accord with rules that are justified by reference to the epistemic goal, which are clear cases of epistemically irrational beliefs. The second is that there is no good reason to think that the epistemic goal is intrinsically valuable.<sup>25</sup>

## 2.4. Summary

To sum up, then: there are a number of philosophers who are not instrumentalists, who nevertheless take instrumentalism seriously, and who offer serious challenges to it. There are also a significant number of epistemologists who adopt the instrumentalist conception.

We have gone through a number of instrumentalists in this chapter. Most of the instrumentalists that we have seen do not take a stand regarding whether they prefer

---

<sup>25</sup> One thing that I think is worth pointing out is that rule-consequentialism is vulnerable to a familiar dilemma: either it collapses into simple act-consequentialism, or else it is internally inconsistent. The idea is that, if the argument for rule-consequentialism is that the adoption of a set of rules will maximize the good, then any rules in that set which forbid particular acts that will maximize the good are not consistent with the argument for adopting rule-consequentialism. That is, if the rule does not permit exceptions in cases where performing an act will achieve more of the good than omitting the act would do, then the rule is not consistent with the goal of maximizing the good. But if the rule permits such exceptions, it will boil down to being extensionally equivalent to act-consequentialism, because the rule will not forbid any action that maximizes the good.

This dilemma is perhaps not a decisive objection to rule-consequentialism, but many have found it to be persuasive. (See Card 2007 and Hooker 2007 for some debate.) The point here is just that it is not obvious that one can avoid the problems that go with act-consequentialism simply by opting for rule-consequentialism. I will not make any more out of this objection here, though, because rule-consequentialism is not the main focus of this dissertation, and because as I just mentioned, some of the arguments in Chapters 6 and 7 will rule out epistemic rule-consequentialism anyway.

subjectivist or objectivist instrumentalism. Some do have explicit positions on that question, though. I will take Richard Foley to be the paradigm subjectivist instrumentalist in what follows, since it is in subjectivist terms that he develops his theory, and his is a very interesting and fully worked-out theory of epistemic rationality. Other important subjectivists are Laudan and Giere.

I will take Alston's earlier work to be the closest thing to a paradigm objectivist instrumental conception of epistemic rationality, because he holds that having a large body of beliefs with a favourable truth-falsity ratio about matters of interest and importance (or some variation on that) is the function of cognition. He is therefore not a subjectivist about the epistemic goal, given that caring about achieving the goal is not relevant to the function of cognition. He also holds that all epistemic evaluation is evaluation relative to the epistemic goal. In his (2005), as we already saw, Alston is no longer interested in epistemic justification or rationality, because he does not think that any single concept could play the role that justification is supposed to play. Still, even though in his later work he is a pluralist about types of positive epistemic status, he is still an objectivist instrumentalist about the various types of epistemic evaluation.

### **Chapter 3: Reasons for the Instrumental Conception**

In Chapter 2, we saw that there are a number of philosophers who take instrumentalism seriously. In this chapter, we will see some possible reasons for wanting to adopt the instrumental conception. Along the way, I will provide arguments for thinking that those reasons are not decisive in favour of the instrumental approach. I do not propose to show conclusively that they are worthless, but I do want to argue that they are not as strong as they appear at first, and that a healthy step back and reassessment of the instrumental conception is in order. The instrumental conception of epistemic rationality, recall, is the view that the epistemic rationality of doxastic states depends on the content of an epistemic goal or goals. Subjective instrumentalism makes the relevant epistemic goal(s) those that agents care about achieving; objective instrumentalism holds that the epistemic goal(s) is (are) valuable independently of what agents care about.

There are a number of reasons which one might give for accepting instrumentalism. I will consider five. First, the instrumental account might be able to offer a deep explanation of why evidence seems to be so important for the determination of the epistemic status of a belief in certain paradigm cases of epistemic rationality and irrationality. Having an explanation of that sort, if one can be given, is certainly more satisfying than simply resting with the claim that evidence just is what determines epistemic status.

Second, an important reason for accepting instrumentalism is found in epistemological naturalism. Naturalism in epistemology is not a unified or well-defined

position, but it is fashionable, and almost everyone wants to travel under the naturalist banner. One common theme shared by many naturalists is that epistemology is not entitled to any kind of normative talk that scientists would reject. Scientists do not reject instrumental rationality or normativity, so perhaps a reasonable way to retain talk of rationality and normativity in a naturalist framework is to adopt an instrumental account.

Third, one might appeal to Bernard Williams-style reasons-internalism in support of the instrumental conception. Reasons-internalism in this sense is the view that in order for a consideration to count as a reason for an agent to do anything, it must bear on the set of desires and attitudes capable of motivating the agent to act. Reasons-internalism provides some motivation for accepting the subjective instrumental conception.

Fourth, one might want to be an instrumentalist simply by default: in order to go about identifying anything as specifically epistemic, one might say, we have to *begin* by looking for what is truth-conducive. It is just natural to think of the truth-goal as epistemically primitive, with everything epistemic deriving from it.

And finally, the fifth reason in support of the instrumental conception is that it offers a way of giving a unified account of epistemic and practical rationality, and unified accounts are better than scattered ones. Instrumental rationality is (at least) a species of practical rationality, so if epistemic rationality is instrumental, then it is a species of practical rationality.

We will go through these potential reasons for accepting the instrumental conception in turn, and along the way, I will provide some reasons for thinking that they

do not give us as good a reason to accept the instrumental conception as it might at first appear. Again, I do not aim to give conclusive reasons for thinking that these reasons do not offer any support for instrumentalism – indeed, I think that some of them at least do have some probative force. But I will try to show that they do not have the force that a defender of the instrumental conception might hope.

### **3.1. A Deep Explanation**

#### **3.1.1. For**

The first reason I want to consider for accepting the instrumental conception is that it might provide an explanation of the force of evidence in certain paradigm cases of epistemic rationality and irrationality. A belief that is held contrary to very strong available evidence is thereby epistemically irrational. It is natural to take that as a datum, which any theory of epistemic rationality ought to accommodate. But despite its being a datum which theories must accommodate, the fact that evidence is so important for the epistemic status of beliefs in paradigm cases is so far an *unexplained* datum. It is all very well to say that a hypochondriac's belief that he has contracted an extremely rare disease, despite the lack of evidence that he has contracted it, is epistemically irrational. But it would be much more satisfying to be able to say *why* it is that evidence is so important in such cases. And indeed, the instrumental conception might be able to offer just the kind of explanation we are after. The idea would be something like the following. The epistemic goal is good to achieve (either because we want to achieve it, or because it is good to achieve even if we do not want to achieve it). And the epistemic goal is (roughly)

achieving the truth and avoiding error. But good evidence for a belief is evidence that makes the belief likely to be true. So to hold beliefs in accord with the evidence is to hold beliefs that are likely to achieve the epistemic goal, and to hold beliefs against the evidence is to hold beliefs that are not likely to achieve the epistemic goal.

Although very few epistemologists offer any explicit arguments of this sort in support of the instrumental conception (with Foley (1993, ch.1) as a notable exception), it is because of the possibility of providing an explanation of this sort that I find instrumentalism to be such an interesting view. Indeed, I would not be entirely unhappy if my main arguments in this dissertation are mistaken, and the instrumental conception turns out to be correct, because it could then provide an explanation of the sort that I would like to have for the importance of evidence.

### **3.1.2. Against**

However, there are two reasons why the prospect of an explanation of the importance of evidence for epistemic evaluations is not a conclusive reason for accepting the instrumental conception. For one thing, this reason is at most a weak reason in support of instrumentalism. It offers an explanation of a phenomenon that we would like to explain, but absent positive reasons for thinking that the instrumental conception of epistemic rationality is correct, the mere fact that it can explain something that we would like to be able to explain does not count for very much. Generally, for a given phenomenon, any number of possible explanations can be offered, in the form of theories which might otherwise be entirely ridiculous, but which could explain the phenomenon perfectly well.

For example: the theory that tiny benevolent gremlins work tirelessly inside my computer to reliably and very quickly produce what I am typing on the screen is a very silly theory, with very little going for it beyond the fact that it can offer an explanation of why the letters that I type always appear almost simultaneously on my screen. The gremlin theory does explain a phenomenon that I find, if not very interesting, at least very important. But before accepting the theory, we need to determine whether it has anything else going for it. (Has anyone seen a gremlin? Is the theory internally consistent? Are there alternative available explanations of the data that we want explained? etc.) Similarly, we need to examine the instrumental conception of epistemic rationality to see how independently plausible it is, before we accept it as an explanation of the importance of evidence in those paradigm cases where evidence is important.

The second reason that this potential explanation of the importance of evidence is not a conclusive reason in support of the instrumental conception is that there might be other explanations to be had. I do not have a fully worked-out alternative explanation to offer, but one might suggest some sort of Kantian explanation, to the effect that believing in accord with the evidence just is part of what constitutes a rational agent, and failing to believe in accord with the evidence is to fail at being a fully autonomous person. Onora O'Neill (1989, ch.1), for example, argues that the categorical imperative, for Kant, is fundamental to both practical and theoretical reason. If that is correct, then a story might be told according to which theoretical rationality (which we might take to either include or else to be identical with epistemic rationality) is a constituent of what it is to be a

rational agent, and (with Korsgaard 2009) we might then argue that in order to be a full person, we must think and act according to the standards of rationality – in particular, with the hypothetical and categorical imperatives. If theoretical rationality is about reasons to think that propositions are true, and evidence is what bears on the truth of propositions, and the categorical imperative underlies theoretical rationality, then we have a categorical explanation of the importance of evidence.<sup>26</sup>

I do not mean to argue that this Kantian line of argument is correct. My aim here is only to illustrate a possible alternative explanation of the importance of evidence. Perhaps this sketch of an account is correct, or perhaps not. Perhaps there is still some other explanation to be had. The important point is that, for the purpose of giving an explanation of the importance of evidence, instrumentalism is not the only game in town.

Also, before moving on, I want to point out that, although a deep explanation that brings together epistemic and practical rationality is certainly desirable (and that is not just an idiosyncratic wish of mine (cf. Heil (1992), Foley (1993, ch.1), and Mills (1998)), it could turn out that epistemic and practical rationality are just different kinds of things, as Feldman (2000) holds. We should not close off this possibility before we evaluate the plausibility of the potential explanations for the importance of evidence for epistemic rationality.

---

<sup>26</sup> Another possible explanation of the force of evidence is that we have a categorical duty of self-respect, and self-respect entails (among other things) believing what we have reason to think is true, i.e. believing in accord with the evidence (cf. Wood 2008, pp. 18-19.) – although this account might turn out to be just an assertion that evidence is important for self-respect, without offering any deep kind of explanation of why that is so.

## 3.2. Naturalism

### 3.2.1. For

Epistemological naturalism is a cluster of views that broadly resemble each other in various ways, most notably in their emphasis on the importance of incorporating empirical, scientific data into our epistemological work. Although there are just about as many variations on naturalism as there are philosophers who think of themselves as naturalists, there is a serious trend today among epistemologists, and among analytic philosophers more generally, to self-identify as naturalists in some important sense.

One way to be a naturalist in epistemology is to hold that epistemologists may not appeal to a priori truths, or to conceptions of rationality or normativity that are not acceptable from a suitable scientific standpoint. Given commitment to a naturalism of this sort, it makes sense to opt for an instrumental account of rationality and normativity, given that means-ends analyses are scientifically respectable.<sup>27</sup>

Quine famously argued that epistemic normativity is not prior to science – that all claims worth investigating about knowledge, evidence, etc., are open to scientific investigation, and that the only kind of epistemology worth doing is empirical research into the way that belief relates to evidence. Here is a widely-cited passage from his “Epistemology Naturalized”:

Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomenon, viz., a

---

<sup>27</sup> I am sliding between talking about rationality and normativity in this section. These are not identical concepts, but they go hand in hand in discussions of naturalism. The reason is clear enough: if, as seems reasonable, rational requirements are a species of normative requirements, then rationality stands or falls with normativity.

physical human subject. This human subject is accorded a certain experimentally controlled input-certain patterns of irradiation in assorted frequencies, for instance-and in the fullness of time the subject delivers as output a description of the three-dimensional external world and its history. The relation between the meager input and the torrential output is a relation that we are prompted to study for somewhat the same reasons that always prompted epistemology; namely, in order to see how evidence related to theory, and in what ways one's theory of nature transcends any available evidence. (1968, pp.82-83)

There is a more or less standard reading of Quine's naturalism, which takes passages like this one at face value. On this reading, Quine wants to get rid of traditional epistemology, together with its concern with normativity and justification, in favour of the psychological study of the way that people in fact go about forming beliefs and theories about the world.

However, Quine claims in some places not to want to get rid of normativity altogether. In another fairly widely cited passage from his "Reply to White," he writes:

Naturalization of epistemology does not jettison the normative and settle for the indiscriminate description of ongoing processes. For me normative epistemology is a branch of engineering. It is the technology of truth-seeking or, in more cautiously epistemic terms, prediction... There is no question here of ultimate value, as in morals; it is a matter of efficacy for an ulterior end, truth or prediction. The normative here, as elsewhere in engineering, becomes descriptive when the terminal parameter has been expressed. (1986, pp.664-665)

So Quine wants to retain normativity in epistemology, and he thinks that the way to do it is to take what is normative to be a function of whatever goals are relevant for epistemology i.e., achieving the truth, or arriving at accurate predictions.

Quine is far from the only philosopher who opts for instrumental accounts of normativity and rationality or justification on the grounds that they are the only accounts

possible in a thoroughly naturalistic framework.<sup>28</sup> However, Quine provides much of the inspiration for naturalistic epistemologists who have come since. Familiar as Quine's naturalism is, it is not necessary to review its various manifestations in the post-Quinean epistemological literature. I would like, however, to cite Elijah Millgram, who clearly articulates a similar line of argument regarding naturalism and instrumentalism about reasons for action (although he is not himself an instrumentalist about practical reasons):

instrumentalism seems to be metaphysically respectable, where the arbiter of respectability in the relevant circles is a broadly-shared image of science. Desires are psychological states, and there is no problem in making room for psychological states among the particles, organisms, causal regularities, and other items that we encounter in science textbooks... And there is in principle no problem in explaining how we come to know what desires are had by whom. The same cannot be said, however, about values, the Good... and other such creatures from the far side of the fact-value distinction. (1997, p.5)

Millgram nicely articulates the naturalist argument for instrumentalism about practical reasons. The naturalist argument for instrumentalism about epistemic reasons is exactly analogous to the naturalist argument for instrumentalism about practical reasons. Desires are scientifically respectable; other alleged sources of rationality are not; desires can ground instrumental rationality; so it is to desires that we must look to ground an account of rationality. The same argument can be given in the cases of both practical and epistemic reasons.

### **3.2.2. Against**

One problem with the appeal to epistemological naturalism to support instrumentalism is that, as Bonjour puts it, "like many fashionable positions, naturalized epistemology is a

---

<sup>28</sup> Others include Giere (1989; 2001), Kornblith (1994; 2002), Kitcher (1992), and Tannsjo (2010).

rather diffuse and uncertain target” (1994, p.283). Very many philosophers want to travel under the naturalist banner. Because it is so common for epistemologists to claim to be naturalists, the appeal to naturalism brings with it a certain feeling of credibility. However, that feeling is not well-grounded; naturalism is less monolithic and solid than it is vague and nebulous. Some types of naturalism do require the adoption of an instrumental conception of rationality and normativity, if they are to permit any rationality or normativity at all, but not all types of naturalism require it, and those that do are themselves far from uncontroversial.

Almeder (1990) helpfully distinguishes three broad types of naturalism in epistemology. The first is *replacement* naturalism. This type of naturalism seeks to replace the traditional kind of epistemology, which focuses on the analysis of knowledge and justification, with the project of “empirically describing and scientifically explaining how our various beliefs originate, endure, deteriorate or grow” (p.263). This is the standard reading of Quine’s naturalism. It seeks completely to replace the analysis of knowledge, the determination of the correct/normative conditions for holding beliefs, etc., with a purely descriptive project. On its face, this first kind of naturalism does not require an instrumental account of rationality or normativity, simply because it does not concern itself with anything normative. However, as we saw, even Quine wants to retain some kind of normative talk. Instrumentalism can retain normativity in such a framework, the idea goes, because the project of describing people’s goals and the best ways to achieve them is scientifically respectable.

The second kind of naturalism that Almeder identifies is *transformative* naturalism. This form of naturalism brings together traditional epistemology and science. It retains a place for the analysis of knowledge and justification in epistemology, but shifts the determination of whether anybody knows anything, and the determination of the cognitive processes at work when we form beliefs, to biology and psychology. Goldman (1986) is a chief proponent of this style of epistemology. This kind of naturalism does not require the adoption of an instrumental account of rationality or normativity; it is perfectly consistent to hold both that justification involves categorical reasons for beliefs, and that whether people ever do hold beliefs in accord with categorically justifying reasons is a matter to be determined by scientific investigation.

Finally, Almeder's third type of naturalized epistemology "simply insists that the method of the natural sciences is the only method for acquiring a proper understanding of the nature of the physical universe" (1990 p.263). Armchair speculations about how the world works are scientifically unacceptable, and they are illegitimate for use in philosophical theorizing. This kind of naturalism is also compatible with categorical reasons and rationality. To learn about how the world works, study the world empirically. But the study of the normative conditions for forming and sustaining beliefs need not be restricted to empirical observations; a priori reflection and argumentation about categorical reasons and rationality are possible in such a framework. Indeed, it should be obvious that the methods of the sciences must themselves be vindicated by rational reflection. For one thing, we cannot simply assume that scientific methods are legitimate.

Some have turned out to be poor methods, and had to be rejected. (For example, think of phrenology, the study of the relation between a skull's shape and size, and the intelligence and personality of the person whose skull it is.) Some scientific methods are not themselves empirical, and require a priori reflection to justify. (Try proving a theorem in pure mathematics by empirical means, and the necessity of a priori reasoning should quickly become obvious.) Far from being contrary to the scientific project, then, a priori reflection is essential to it.

So, of the three types of naturalistic epistemology that Almeder identifies, only one requires the adoption of an instrumental conception of rationality. What is more, there are versions of naturalism that do not fall into these categories, that also do not require the adoption of an instrumental approach to rationality. For example, consider William Alston's naturalism. Alston takes the attempt to respond to the radical skeptic about knowledge and justification to be the attempt to do epistemology as first philosophy, and he holds that that is the wrong way to do epistemology. His naturalism consists of the refusal to take the skeptic seriously. He talks about naturalism as follows:

...not the extreme version put on the map, unfortunately, by Quine... but a more moderate version that is distinguished precisely by avoiding the temptation to play the skeptic's game. One declines to pursue epistemology as "first philosophy," an attempt to get conclusions as to what we know or how we know before we address ourselves to getting any knowledge about anything else. Instead, one approaches epistemology in the same "natural" spirit as any other problem area – by working with any of our knowledge, beliefs, or assumptions that seem to be of relevance to the problems at hand; remembering, of course, that any of them can be called into question at a further stage of inquiry. (2005, p.8)

...from a "naturalistic" standpoint... we feel free to assume relevant things we take ourselves to know or well-groundedly believe... (ibid., p.201)

...the “naturalistic” approach to epistemic evaluation... involves making use of whatever we take ourselves to know, or believe on an adequate basis, when we epistemically evaluate a belief or a class of beliefs. (ibid., p.191)

What is important here is that Alston’s brand of naturalism consists only in the rejection of the skeptical game and the attempt to do first philosophy. And, although Alston is an instrumentalist about epistemic evaluation, his instrumentalism is not required by his naturalism. One can simultaneously both refuse to respond to the skeptic and hold that epistemic reasons are categorical.

Or consider Alvin Plantinga’s naturalism. Plantinga claims to be a naturalist, in the sense that he only wants to allow himself the kind of normative talk that scientists allow themselves (1993, pp.45-46; 210-211). He famously grounds warrant, his preferred epistemic term of appraisal, in the proper function of truth-directed cognitive systems. Proper functions are widely held to be scientifically respectable, so appealing to proper functions fits with this kind of epistemological naturalism. (We will see more on proper functions in Chapter 5.)

However, Plantinga also argues that epistemological naturalism, as well as any kind of scientific enterprise that makes use of the notion of a proper function, is at home only in a metaphysical *supernaturalism* (1993, ch.11). The idea is that in a hard-nosed metaphysical naturalism, there is no ultimate source of design or normativity, so if we want to talk about proper functions (which many of us do), we have to embed such talk in a broader supernatural metaphysics.

Now, granted that epistemological naturalism need not go hand in hand with metaphysical naturalism, still the two naturalisms usually do go together, and it is easy to see why: if one endorses a metaphysics that incorporates supernatural elements, then there is little reason to restrict one's epistemology to elements that are respectable from the point of view of the (natural) sciences. On the other hand, if one endorses a naturalist metaphysics, involving such claims as that humans are part of the natural order, and that the things in the space-time continuum are all that there are, then one will naturally gravitate toward a naturalist epistemology (cf. Brown 1988). Plantinga's attempt to bring together epistemological naturalism with metaphysical supernaturalism seems, at the very least, to be a bit forced. But, more importantly, if we drop the commitment to metaphysical naturalism, then there is no reason to restrict ourselves to the kind of normative claims that are considered respectable by natural scientists, because what natural scientists study does not exhaust what there is.

Finally, some philosophers want to bring together metaphysical naturalism and categorical rationality. Carrie Jenkins has recently made an interesting attempt in that spirit:

Naturalism, for the purposes of this paper, is the position which holds that some scientific world view is approximately correct, so that there exist no supernatural or otherwise spooky entities, properties, events or other phenomena. If naturalism is correct, then the world is the way that best science says it is (where best science need not be identical to any actual science, whether present or future), and there is nothing more to be truly said about the world. (2007, p.259)

(Notice, by the way, that this characterization of naturalism is largely metaphysical.) A naturalist like Quine would no doubt agree with this characterization of naturalism, and

would take this kind of naturalism to require an instrumental conception of normativity, if there is any normativity at all. Jenkins's strategy for avoiding that consequence is to identify normative facts with natural facts:

The strategy to be defended here is supposed to be a way of maintaining that epistemic normativity is a genuine, distinctive and non-instrumental kind of normativity, that claims of epistemic normativity are true, and that they are made true by facts in the world. The idea is that they are made true by the same facts as certain other claims which are more obviously naturalistically acceptable... On the view under consideration, the epistemically normative facts are not eliminated but simply identified as certain natural facts. (ibid, p.261)

I have no settled opinion about Jenkins' position. However, what is important is that we have here an attempt to get a non-instrumental account of epistemic normativity off the ground within a thoroughgoing metaphysical naturalism. If that attempt can be made to work, then even a hard-nosed naturalism that denies the existence of "spooky" non-natural entities will not *require* the adoption of an instrumental conception of rationality or normativity.

I do not take myself to have refuted naturalism in epistemology. The kind of objection I do take myself to have defused is the following: "Look, *any* clear-headed person has to accept naturalism. And naturalism entails instrumentalism. So if you want to be a serious epistemologist, you have to be an instrumentalist." The point of this section was to chip away at the naturalist motivation for the instrumental conception by way of showing that there are a number of ways to be a naturalist without adopting an instrumental conception of epistemic rationality. The ways to be a naturalist are many and varied. And the type of naturalism that *does* require instrumentalism is not nearly as

widespread a view as one might think, given how common it is to appeal to naturalism nowadays.

### **3.3. Reasons-Internalism**

#### **3.3.1. For**

The third reason for accepting instrumentalism about epistemic rationality is internalism about reasons, as it is understood in metaethics and action theory. The idea is that a reason for an agent to  $\phi$  necessarily involves a pro-attitude about  $\phi$ -ing. Davidson (1963) famously held that reasons for action consist of a pro-attitude of some sort about  $\phi$ -ing, together with some beliefs about  $\phi$ -ing. It is because of their belief+attitude structure that reasons can explain actions. Bernard Williams (1981) similarly held that a reason for an agent to  $\phi$  must make reference to her “subjective motivational set,” i.e. the set of desires and attitudes capable of motivating her to act.

If one is an internalist in this sense, then one might be tempted to opt for an instrumental account of reasons. The idea is that if we take an agent’s set of desires, values, etc. (her subjective motivational set) to be identical with the set of her goals, then because all reasons must make reference to her subjective motivational set, it follows that she can only have reasons to perform actions that promote the achievement of her goals. In his argument for an error theory about morality, for example, Richard Joyce (2001) employed reasons-internalism in order to reject categorical conceptions of morality. Millgram puts the reasons-internalist argument for practical instrumentalism clearly and succinctly: “How could anything be a reason for action if it could not motivate you to

actually *do* something? And what could motivate you to do something except one of your desires?” (1997, p.4).

This kind of argument is not widely employed in support of the instrumental conception of epistemic rationality. However, Brian Huss at least offers an argument of this kind in support of a type of instrumental conception of epistemic reasons<sup>29</sup>:

[Theories] based on categorical imperatives can lack the force to motivate people to think or act in the way that is said to be rational. Hypothetical Imperatives, on the other hand, do not seem to have this problem. If you invoke a categorical imperative and tell me I should do x, I may be inclined to ask, “Why should I?” But if you use a hypothetical imperative and tell me I should do x because I want y and because x is the best means to y, then it is difficult to even understand the why-should-I question, and it seems people would be unlikely to ask it. (2009, p.252)

Although Huss writes here in terms of *thinking* and *doing* x, he is defending an instrumental conception of reasons for belief.

There is an interesting argument, then, in support of the instrumental conception of epistemic rationality, which takes its support from motivational internalism about reasons, and I want to address this argument.

### **3.3.2. Against**

There are two things to say about this argument for instrumentalism.<sup>30</sup> The first is that, even if it works, it only works as an argument for the subjective instrumental conception

---

<sup>29</sup> Huss allows that there may be a sense of “epistemic reason” that is not instrumental, although he does not take that to be a very interesting category of reasons.

<sup>30</sup> In fact, there is a third thing to say: the subjective motivational set that is typically appealed to is not the set of what can motivate the agent as she is in the actual world, but rather, it is the set of what *would* motivate the agent to act, *if* she were fully informed and ideally procedurally rational. This issue is somewhat beside the point for our purpose here, but I want to point out that it is very strange, when we are concerned with what can motivate an agent to do something, to focus on what would motivate a fully

of epistemic rationality, because that is the view according to which the epistemic goals that people care about achieving are what determines their epistemic rationality.

Proponents of the objective instrumental conception will have to look elsewhere to support their view, given that they appeal to epistemic goals that agents need not care about (and which therefore need not be capable of motivating them).

The second thing to say is that there is an important distinction to be drawn between the different levels at which desires can be appealed to in an account of reasons. We can best get at the distinction by considering Searle's (2001, ch.6) argument against motivational internalism about reasons, and a natural objection to Searle's argument. Searle argues that agents are capable of giving themselves reasons for action, which will continue to be reasons for them to act at a later time even if at that later time they have no desire that the action will satisfy. We can *commit* ourselves to courses of action; our commitments can outlast our desires to fulfill them, and we can have reasons to act on our commitments. Searle's example is that you can have a reason to pay for a beer at a bar, after you've finished drinking it, even if doing so will not fulfill any of your desires. (Imagine that the bar is in a town you'll never come back to; the bartender is busy and won't notice you leave; there is no bouncer; and so on, so that there are no desires that you have that paying for your beer will satisfy.) The only kind of reason that there could be to act is a non-instrumental one. One might call it a moral reason.

---

informed and ideally procedurally rational version of an agent to act. I find that so strange, because agents typically are *not* either fully informed or ideally procedurally rational, and what could motivate ideal versions of ourselves is not what could motivate us as we really are. The appeal to this kind of motivation therefore seems unmotivated.

A natural objection to Searle's argument is that in order for non-instrumental moral reasons to get you to pay for your beer in the example above, you must still *want* to be a moral person, or want to be the kind of person who fulfills her commitments, or something like that. So there is still a sense of reasons-internalism that is true: there must be some perceived desire-satisfaction to be obtained by performing the action in order to be motivated to perform it.

That objection is correct, and it points to the distinction that we need to draw: we need to distinguish different locations for the desires that can motivate us to act. In ordinary, simple cases, reasons for action consist of beliefs and desires: the desire to be satisfied, and the belief that the action in question will satisfy the desire. But there can be cases that are less simple, where an agent can have (for example) a moral reason to perform an action, and the agent can be motivated to perform the action, even though the action does not promote the satisfaction of any of the agent's desires *other* than the desire to do what is morally required (whatever that may turn out to be). In such cases, what is morally required need not *itself* depend on any of the agent's desires; it can motivate the agent to act by virtue of her desire to do what is morally required.

Perhaps an example will help here. If you promise to take your cousin to the zoo, but then some tickets for your favourite show fall into your lap, you can have a moral reason to take your cousin to the zoo, even if you have no (non-moral) desire that doing so will satisfy. (Suppose that you do not much like your cousin, so you do not care about whether he will be less happy if he does not get to go to the zoo; no one will find out if

you back out on your promise; etc.) You can still recognize that there is a moral reason for you to fulfill your promise. Now, perhaps you are the kind of person who wants to be moral, or perhaps not. If you do not care about being a moral person, then the mere fact that you had made the promise could not motivate you to fulfill your promise (provided that you had no other desire that fulfilling your promise would satisfy). If you do want to be moral, on the other hand, then you will be capable of fulfilling your promise just because you are morally required to do so. The point here is that the moral requirement itself does not depend on the existence of any of your desires; you have the desire to take your cousin to the zoo precisely because that is what is morally required.

Even if there is a sense of reasons-internalism that is true, then, it is consistent with a qualified type of reasons-externalism. To put the point more explicitly: if motivational internalism about reasons is correct, then in order for an agent A to have a reason to  $\phi$ , it must appear to A as though  $\phi$ -ing will do well vis-à-vis her subjective motivational set (i.e., her desires, broadly construed). But the desire to be moral is itself a desire that can give A a reason to  $\phi$ , if  $\phi$ -ing is what morality requires – *and the moral requirement to  $\phi$  need not itself depend on any of A's desires.*

Similarly, there can be epistemic reasons that do not themselves depend on anyone's desires. It may very well be that in order to form beliefs in response to epistemic reasons, there must be a desire that doing so will satisfy. But that internalist requirement does not entail that the epistemic reason itself must depend on the existence of a desire to achieve an epistemic goal. In order for an agent to have an epistemic reason, in the sense

of being motivated to believe in accordance with it, it is sufficient that she have the desire to believe in accord with her epistemic reasons.

And – this point is important – the mere desire to believe in accord with epistemic reasons cannot itself be the desire that acts as an epistemic goal, as what grounds epistemic rationality, in the way required by the instrumental conception. According to the instrumental conception, there is (are) some epistemic goal(s), which is (are) promoted by holding certain beliefs. Those beliefs are epistemically rational, on this account. But if the instrumental account holds that holding beliefs in accord with one’s epistemic reasons can be the epistemic goal, then the account of epistemic rationality becomes circular (epistemic reasons themselves being part of, or at least depending on, the account of epistemic rationality).

### **3.4. The Default**

#### **3.4.1. For**

The fourth reason for accepting instrumentalism is that having a truth-goal in mind is just the way that we have to begin identifying the domain of the epistemic. Truth is important; a truth-centered epistemology is extremely plausible; and the natural way to have a truth-centered epistemology is (allegedly) to begin with a truth-centered epistemic goal.

William Alston, for example, writes:

I don’t know how to prove that *the acquisition, retention, and use of true beliefs about matters of interest and/or importance* is the most basic and most central goal of cognition. I don’t know anything that is more obvious from which it could be derived. But I suggest that anyone can see its obviousness by reflecting on what would happen to human life if we were without beliefs at all or if our beliefs were all or mostly false. (2005, p.30)

Jarrett Leplin expresses a similar point:

I do not attribute what I take to be the epistemic goal to individual cognizers, nor do I claim to read it off of epistemic practice, say as the best explanation of what cognitive agents do. What agents do underdetermines their goals. Real agents are many things besides cognizers, and I would not know how to identify the cognitive part of practice without an epistemic goal already in mind. I simply assume that believing truly has intrinsic value, and that this value is codified in a goal that is distinctively epistemic, as against, say, moral, aesthetic, or pragmatic. (2009, p.19)

The idea is that we demarcate the realm of the epistemic by beginning with an epistemic goal already in mind, and it's just obvious that truth is constitutive of that goal.<sup>31</sup>

### **3.4.2. Against**

The response to this reason for instrumentalism is that truth can play a central role in the account of epistemic rationality without making the account instrumentalist. Take a traditional categorical evidentialism, for example. A typical evidentialist claim is that one ought to believe that  $p$  only when there is sufficient evidence that supports  $p$ . But evidence for a proposition is nothing other than whatever bears on the truth of a proposition, so traditional evidentialism has truth playing a central role. The natural thing to think *may* be that if we are to have a truth-centered epistemology, truth must be a goal toward which our practices are directed – but the natural thing to think is not always

---

<sup>31</sup> A related way to argue for instrumentalism, together with a truth-centered epistemic goal, is to point out that knowledge essentially involves non-accidental truth, so we have to begin epistemological theorizing with achieving truth and avoiding error as a goal. Paul Moser writes, for example, that “Any standard or strategy worthy of the title “epistemic” must have as its fundamental goal the acquisition of truth and the avoidance of error. This follows from the fact that genuine knowledge has truth as an essential condition and excludes error” (2002, p. 14). However, as we are concerned here with rationality or justification, rather than knowledge, we do not need to consider this way of motivating instrumentalism.

correct, and there are non-instrumentalist ways to make truth play a central role in an account of epistemic rationality.

### **3.5. A Unified Account of Rationality**

#### **3.5.1. For**

The fifth supporting reason for the instrumental account of epistemic rationality is that it allows us to give a unified account of epistemic and practical rationality. Unified accounts are elegant, and to be preferred over messy, disjunctive accounts of rationality. As we saw in Chapter 2, for example, Richard Foley's view is that epistemic rationality is a variation on the "general template of rationality," which is: "an action A (or decision, plan, intention, strategy, belief, etc.) is rational for a subject S if it is rational for S to believe that A would acceptably satisfy her goals" (2008, p.45). The general template provides a framework within which to assess rationality of any kind. Epistemic rationality, too, is an instance of the template.

Perhaps a stronger way to put the point is, as Larry Laudan puts it, that "there is no coherent sense of justification (epistemic or otherwise), just as there is no sense of deliberative action (epistemic or otherwise), in the absence of the specification of the ends with respect to which an action is deemed justified or rational" (1990, p.317). Laudan's view *requires* that there be a unified analysis of epistemic and practical rationality, because he holds that there is only one concept of rationality that is even *coherent*.

### 3.5.2. Against

First of all: it is possible that instrumental rationality does not exhaust practical rationality; perhaps there is more to practical rationality than instrumental concerns. If so, then the desire to give a unified account of epistemic and practical rationality will not require that we have an instrumental account of epistemic rationality. But I will not push that point here, since I do not have a settled view of practical rationality, and arguing for such a position would be well beyond the scope of this dissertation anyway.

If practical rationality *is* in fact exhausted by instrumental rationality, then there is little to say to undermine this motivation for the instrumental conception. Unified accounts are more elegant and are to be preferred, other things being equal, to messier accounts. However: (1) this reason for accepting instrumentalism is only a rather weak reason; if we end up with an account of epistemic rationality that does not have a similar structure as the best account of practical rationality, that might be a price worth paying. So even though there is little to say to *undermine* this reason for instrumentalism, there might be plenty to *override* it. And (2) although unified accounts are preferable to messier accounts, other things being equal, much of the point of this dissertation is to show that other things are *not* equal, because the instrumental conception faces some serious challenges.

In response to Laudan's stronger way of putting the point – that instrumental rationality is the only coherent type of rationality – again, there is little to say except that there do appear to be accounts of rationality that are non-instrumental, which are not

thereby incoherent. Think, for example, of Kant's rigorous and not apparently incoherent categorical system of theoretical and practical reason. One might of course say that Kant's view is mistaken, or that it is empty, but it does not appear to be incoherent.

### **3.6. Summary**

In this chapter, we saw five possible reasons for accepting the instrumental conception of epistemic rationality: giving a deep explanation of the importance of evidence for the rationality of beliefs, epistemic naturalism, reasons-internalism, having a truth-goal as a default starting-point for epistemology, and giving a unified account of epistemic and practical rationality. The aim here was not to argue directly against the instrumental conception, but rather to undercut some arguments that might be given in its favour. Some of these arguments have something to be said for them, but I hope to have shown that none of them requires that we accept the instrumental conception, or at least to have shown that a healthy step back and re-evaluation of the instrumental conception is called for.

## **Chapter 4: The Epistemic Goal**

### **4.1. Some Preliminaries**

In the second chapter, we saw that there are a number of people who take the instrumentalist approach seriously. In the third chapter, we saw that there appear to be good reasons to take it seriously, though we also saw that those reasons are not as compelling as it seems at first. According to the instrumental conception of epistemic rationality, again, the epistemic rationality of beliefs (or doxastic states generally) depends on the content of the relevant epistemic goal or set of epistemic goals. The subjectivist version of the instrumental conception identifies the relevant goal(s) as the goal(s) that epistemic agents care about achieving. The objectivist version holds that the relevant goal or goals are independent of what epistemic agents care about achieving.

The present chapter undertakes to determine the most plausible formulation of the epistemic goal(s), because an implausible formulation of the goal(s) will give rise to some obvious and avoidable objections to instrumentalism. A plausible formulation of the epistemic goal(s) will give the instrumentalist position the best reading we can give it. The account of the goal given here begins by drawing a distinction that is not drawn in the literature, as far as I can tell, but it is an important one to draw. Briefly, the distinction is between what I am calling epistemic goals and epistemic values. The idea is that there are a number of things that are valuable from the standpoint of epistemology, but only some of those valuable things get to count as those things in virtue of which doxastic attitudes count as epistemically rational or irrational. Whatever is valuable from the

standpoint of epistemology is what I am calling epistemic *values*; those epistemic values that determine the status of beliefs as epistemically rational or irrational, I am calling epistemic *goals*. (We will get into this distinction more fully in section 4.2.) Drawing this distinction will help instrumentalists in their attempt to get the extension of epistemically rational beliefs correct; and it will make room for monism about epistemic goals while admitting pluralism about epistemic values.

We do find similar distinctions in epistemology, in (Kvanvig 2005), and in practical philosophy and argumentation theory, in (Atkinson et. al. 2004; 2006). Atkinson et. al. draw a helpful distinction between goals and values in practical reasoning, but theirs is importantly different from the distinction advanced here. Their distinction is between particular goals that are desired, and the values that one seeks to promote by achieving those goals. Goals are not values, on their view; they are different kinds of things, existing at different levels, with values at the basic level, and goals depending on values. By contrast, the distinction I am advancing here has it that goals are one kind of value, the kind of value in virtue of which beliefs and other doxastic attitudes count as epistemically rational or irrational.<sup>32</sup>

Kvanvig's distinction is between goals considered from the perspective of the theoretician, on the one hand, and goals considered from the perspective of the subjects whom the theoretician is studying, on the other. Although this distinction resembles mine in some ways, and although I endorse some of its motivation ("Asking about goals is

---

<sup>32</sup> Atkinson et. al.'s distinction is not a rival to mine; I take it that we are simply offering different stipulative theoretical definitions. These definitions are related to the ordinary ways in which the terms are used, but are not intended to be accurate analyses of the ordinary senses of the terms.

asking about values or goods” (2005, p.285)), it differs in two crucial respects. First, Kvanvig is concerned to point out that the goals of the subject and the goals of the theoretician can diverge; they might have little or nothing in common. But epistemic goals, as I am understanding them, are a subset of epistemic values. So, although the sets of epistemic goals and epistemic values need not be identical, they cannot come apart completely. The other respect in which our distinctions differ is that, as I am understanding them, epistemic goals are those values in virtue of which beliefs count as epistemically rational or irrational – epistemic goals are the crucial epistemic values for the purpose of making epistemic evaluations. By contrast, Kvanvig holds that the important epistemic goals/values for making epistemic evaluations are those of the theoretician.

With those preliminary points in mind, we can move on to begin formulating the epistemic goal(s). This chapter addresses three questions about how, assuming that there is such a thing, to formulate the epistemic goal. First, we need to determine what values to include in the formulation of the epistemic goal: is truth-achievement (and error-avoidance) the only value to include, or should we include other values, such as simplicity, coherence, and so on? Second, is the epistemic goal solely about matters of interest and importance, or do unimportant truths count too? And third, is the epistemic goal synchronic or diachronic? That is, is the goal (for example) to believe truths and avoid errors *right now*, or is it to believe truths and avoid errors *over the long run*?

Although I will propose answers to these questions, and set out what I take to be the most plausible formulation of the epistemic goal, I will not be wedded to very much of what follows in this chapter. My goal here is to set out the most plausible formulation of the epistemic goal in order to give the instrumental conception the best hearing possible; a poor formulation of the goal will give rise to avoidable objections to the instrumental account. So I will consider possible answers to the three questions listed above, in order to show the instrumental conception in its best possible light, so that we can later see that there are still unavoidable objections to the instrumental conception.

I cannot say that there are no other interesting questions to ask about the epistemic goal(s), but these three are the ones that have tended to be asked in the literature. Moreover, it is clear why we should ask these questions. The opposing answers to these questions each have some appealing features, either because they are independently plausible views, or because they will help to avoid getting the extension of epistemically rational and irrational beliefs wrong. For example, take the second question. It is clear that matters of interest and importance can be our goals, but what about matters that are of neither interest nor importance? That is not so clear. (I will discuss what it means for a proposition to be of interest or importance in section 4.3.) It is therefore reasonable to restrict the epistemic goal to matters of interest and importance. On the other hand, there can be beliefs about matters that are of neither interest nor importance, that are nevertheless obviously epistemically rational or irrational, so if epistemic rationality

depends on the epistemic goal, perhaps we ought not to restrict the goal with an “interest and importance” clause.

Similarly with the other two questions: one of the possible answers is motivated by reflection on the nature of goals, while the other answer is motivated by the requirement that we get the extension of epistemically rational and irrational beliefs correct. The three questions are tackled in turn in what follows.

## **4.2. Epistemic Goals and Values**

### **4.2.1. A first pass**

In the literature on epistemic goals and values, one main divide is over whether the primary epistemic goal is truth-directed, or whether there are other epistemic goals that are not reducible to the truth-goal. In an oft-cited passage from “The Will to Believe,” William James famously wrote that our “duty in the matter of opinion” is to “*know the truth*” and to “*avoid error*,” and that “these are our first and greatest commands as would-be knowers” (1949/1896, p.99, emphasis in original).

James puts the point in terms of duties as would-be knowers, but many epistemologists pick up James’s emphasis on achieving truth and avoiding error in their formulations of the epistemic goal. For example, William Alston (1985, p.59) holds that the “epistemic point of view... is defined by the aim at maximizing truth and minimizing falsity in a large body of beliefs,” and he writes a little further on that “our central cognitive aim is to amass a large body of beliefs with a favorable truth-falsity ratio.” More recently, Alston has revised his formulation of the epistemic goal, so that it includes

pragmatic considerations: his position is now that the epistemic goal is “maximizing true beliefs and minimizing false beliefs about matters of interest and importance” (2005, p.32). Richard Foley holds that the epistemic goal is “now to believe those propositions that are true and now not to believe those propositions that are false” (1987, p.8). Marian David loosely characterizes the epistemic goal, which he calls the “*truth-goal*,” as “the goal of believing truths and not believing falsehoods” (2001, p.152, author’s emphasis). Keith Lehrer’s view is that the “goal of justification” is “accepting something if and only if it is true” (1990, p.82). Andrew Latus (2000, p.31) holds that the epistemic goal is either to “amass a large body of true beliefs without also holding too large a body of false beliefs,” or “to believe all the truths there are and only those truths.” Wayne Riggs is explicit about wanting to follow James; he takes the epistemic goals to be “having true beliefs and avoiding error” (2008, p.6).

Many theorists, then, tend to accept truth-monism, the view that there is a single epistemic goal, and that that goal is truth-directed. I should point out that ‘truth-monism’ might be a bit of a misnomer: no one takes achieving the truth to be all that there is to the epistemic goal. Avoiding error is certainly important to the epistemic goal as well. It is not clear whether it is better to think of these as separate goals to be achieved, or whether they are best thought of as two parts of a single goal. We might want to take them to be two parts of a single goal, since in the typical case at least, having a true belief about  $p$  rules out having a false belief about  $p$ , and having a false belief about  $p$  typically rules out having a true belief about  $p$ . (I say the *typical* case, because it might be possible to have

contradictory beliefs about  $p$ , in which case having the belief that  $p$  does not rule out the belief that not- $p$ . But such cases are, even if possible, certainly not normal.). On the other hand, as Riggs emphasizes, we can value achieving the truth far more than avoiding error (or vice versa), so perhaps it is best to take achieving the truth and avoiding error to be separate goals that we weigh against each other.

I will not take a stand on whether truth-achievement and error-avoidance are separate goals, or two parts of a single goal. However we think of the relation between achieving the truth and avoiding error, the point is that those who are often thought of as truth-monists want both of them, and nothing else, to count in our epistemic evaluations. That is what pluralists deny: they hold that there are other epistemic goods that are not reducible to the value of truth (and error-avoidance). Kvanvig (2005, p.287), for example, holds that there is a range of epistemic goals, including “knowledge, understanding, wisdom, rationality, justification, sense-making, and empirically adequate theories.” The contrast, then, is between conceptions of the epistemic goal as consisting solely of achieving the truth and avoiding error, and conceptions of the epistemic goal as consisting of more than that. For ease of exposition, I will continue to talk about achieving true beliefs and avoiding false beliefs under the rubric of a monistic truth-goal.

Hess (2010) has recently claimed that monists and pluralists mostly agree that there are in fact a number of epistemic goals, but they disagree over whether the goals are all equally valuable. Hess’s claim is not quite right, though: there is certainly widespread agreement that truth is not the only thing that is valuable from the epistemic standpoint,

but what makes the difference between a monist and a pluralist is not whether the various values are *equally* valuable. Rather, the question is whether the other values are *reducible* to the value of truth or not. Sosa, for example, writes: “Our worry requires only that we consider truth the fundamental epistemic value, the ultimate explainer of other distinctively epistemic values” (2007, p.72). The question is one of reducibility, or explainability in terms of the fundamental value, rather than one of relative degrees of value. A truth-monist can hold that, say, justification is valuable, but only as a means to truth, and still be a truth-monist (BonJour (1985) does this). Indeed, one might go further: one might hold that justification is only valuable because truth is valuable, and so be a truth-monist, but also hold that because we are better positioned to assess whether beliefs are justified or unjustified than we are to assess whether beliefs are true or false, justification is equally as valuable to us as, or even more valuable to us than, truth.

#### **4.2.2. Epistemic goals vs. epistemic values**

Before weighing in on the debate between monists and pluralists, though, we have to draw a distinction between epistemic goals and epistemic values. Epistemic *values* are whatever is valuable from the standpoint of epistemology. (We can provisionally take the standpoint of epistemology to consist of the primary targets of epistemologists’ theorizing.) Epistemic *goals* are those epistemic values by virtue of which, according to the instrumental conception of epistemic rationality, beliefs count as epistemically rational or irrational.<sup>33</sup>

---

<sup>33</sup> In a sense, talking about epistemic goals as a subset of epistemic values seems to beg the question in

Since this is not a distinction that has been clearly drawn yet, epistemologists have often not been explicit about whether they are talking about goals or values, in my sense of the terms, when they talk about goals and values. Most appear to be interested in the question of epistemic values (in my sense), especially those who espouse pluralism. Many are simultaneously interested in the question of epistemic goals, though.

Admittedly, it is not immediately obvious why we should want to draw the distinction between goals and values – after all, if something is valuable from the standpoint of epistemology, why should we not adopt it as an epistemic goal? Why, for that matter, should the rationality of our beliefs not be weighed in terms of it? That line of reasoning is no doubt a major reason why this distinction has not been drawn.

There are two reasons to make the distinction, one minor and one major. The less important reason is in order to be charitable to the instrumental conception of epistemic rationality. In order for that conception to be plausible, the formulation of the epistemic goal must be such as to get the extension of epistemically rational and irrational beliefs right. But if we allow beliefs held in such a way as to promote the achievement of epistemic values other than truth to thereby count as epistemically rational, it is possible that we will have clear cases of epistemically irrational beliefs that the instrumentalist will be committed to counting as epistemically rational, and vice versa, in cases where the

---

favour of the instrumental conception, because it seems to imply that the epistemic goal is a valuable thing to achieve. We will see some arguments for thinking that the epistemic goal is not always valuable to achieve, in Chapter 7, but for now, I am only trying to get the formulation of the epistemic goal down as charitably as possible, so for the moment I am content to play the instrumentalist's game.

goal of truth and the other epistemic goals come apart. Consider, for example, the following case.

**Wendy**

Wendy is a taxonomist who comes up with a brilliant scheme for classifying insects. Her scheme works explanatory wonders, and it enjoys a very high degree of coherence both internally and in relation to other sciences. However, it has one oddity: it entails that ants have only five legs.

Should Wendy therefore believe that ants have five legs? Obviously not. Ants can be observed to have six legs. But the belief is entailed by a highly coherent explanatory system. This case is constructed to be an obvious case where evidence for a proposition's truth and its coherence with an explanatory system pull in opposite directions. If we were to include coherence as an epistemic goal, we would have to say that believing that ants have five legs has something to be said for it. But it has nothing to be said for it, since it is demonstrably false.

That case is an odd, artificial one, but there are real cases that illustrate the point, too. Consider the common assertion that bumblebees *shouldn't* be able to fly, but – amazingly! – they defy our science, and they fly. That is another case where coherence with a scientific theory pulls apart from obvious truth, and truth has to win out. The fact that a scientific theory entails that bumblebees cannot fly does not give us any reason to think that they really cannot fly, when we are capable of directly observing them in flight.

The fact that bumblebees can be observed to fly means that the theory according to which they cannot fly is at least partly mistaken.

Perhaps the pluralist instrumentalist can try to avoid this consequence, by holding that when the goal of truth and the other epistemic goals come apart, the goal of truth is always the more important goal, epistemically speaking. However, in order for the pluralist thesis about epistemic goals to be interesting, it must allow that the goal of truth and the other epistemic goals can come apart, and that truth does not always outweigh the other epistemic goals when they do come apart. Otherwise, the pluralist thesis is so watered down as to be indistinguishable from monism.

It appears, then, that the pluralist account of epistemic goals fails to yield the correct verdict regarding the epistemic rationality of beliefs in certain cases. That is the first reason to distinguish epistemic goals from epistemic values: some epistemic values, such as coherence, when treated as goals by reference to which beliefs count as epistemically rational or irrational, generate the wrong result about the epistemic status of some beliefs.

The other, more important, reason to distinguish epistemic goals and epistemic values is that knowledge and justification are epistemically valuable if anything is, but they cannot be epistemic goals, on pain of circularity. Thousands of pages' worth of ink has been spilled in the attempts to get clear on knowledge and justification; to deny that they are valuable from the epistemic standpoint would be simply incredible. The account of epistemic value ought therefore to have a place for the value of knowledge and

justification.<sup>34</sup> However, because epistemic goals are those goals relative to which, according to the instrumental conception, evaluations of epistemic justification are made, it would be circular for the instrumentalist to hold that epistemic justification can *itself* be an epistemic goal. Similarly, because knowledge involves justification as a constituent (on the traditional account of knowledge, at least), knowledge cannot be a goal by virtue of which beliefs count as epistemically justified.<sup>35</sup> The same will apply for any other epistemically valuable state or property that involves justification as a constituent.

So there has to be a distinction between those epistemic values to which instrumentalism can appeal in its account of epistemic justification, on the one hand, and those epistemic values that cannot enter into the account of epistemic justification, on the other. With this distinction in hand, we can tackle the question of the content of the epistemic goal.

#### **4.2.3. Pluralism vs. monism**

It seems to me that the most plausible take on the epistemic goal is truth-monism: the goal is to achieve true beliefs and to avoid false beliefs (subject to qualifications to be made in the following sections).<sup>36</sup>

---

<sup>34</sup> To be clear, I am not addressing what has recently come to be called the “value problem” in epistemology (cf. Pritchard (2007), Kvanvig (2003)). In one form, that is the problem of whether knowledge is valuable over and above the value of true belief. In another form, it is the problem of whether knowledge has any value over and above any proper subset of its parts – i.e., whether meeting a suitable Gettier-defeating condition adds any value to a justified true belief. I am only concerned here to point out that knowledge and justification *are* epistemically valuable; their relative values when compared to each other or to other values is a separate question.

<sup>35</sup> David (2001, p. 154) makes a similar argument.

<sup>36</sup> Bear in mind that in this section, I do not mean to take any strong stands about the value of truth; I am only trying to give as plausible a formulation of the epistemic goal as I can, in the interest of giving the instrumental conception its best possible hearing.

There are a few avenues by which truth-monism gains its support. The first is the fact that, as we have just seen, we are forced to draw a distinction between epistemic goals and values, so it does not follow merely from the fact that something is epistemically valuable, that it is also an epistemic goal. That is, we have shown that there must be a species of epistemic values that are not epistemic goals, including both knowledge and justification; it is reasonable to think that some other epistemic values might fall into that category, especially given that truth-monism works so well otherwise.

The second avenue of support for truth-monism is from the nature of beliefs, and the kinds of things that are natural to think might count as reasons for beliefs. The concept of belief is a somewhat vexed one, and I for one do not have a clear, developed set of individually necessary and jointly sufficient conditions for belief. But I think that there are nevertheless informative things to say about belief that will help to support truth-monism with respect to the epistemic goal. First of all, it is a commonplace in contemporary analytic philosophy to remark that to believe a proposition is to take it to be true.<sup>37</sup> This commonplace, if correct, yields a plausible definition of “belief”: for all subjects  $S$  and all propositions  $p$ ,  $S$  believes that  $p$  iff  $S$  takes  $p$  to be true, and  $S$  disbelieves that  $p$  iff  $S$  thinks that  $p$  is false.

It is not easy to say just what is involved in taking  $p$  to be true, however. If we want to think of taking  $p$  to be true as what it is to believe that  $p$ , then taking  $p$  to be true

---

<sup>37</sup> Eric Schwitzgebel’s entry “Belief,” in the *Stanford Encyclopedia of Philosophy*, for example, begins: “Contemporary analytic philosophers of mind generally use the term “belief” to refer to the attitude we have, roughly, whenever we take something to be the case or regard it as true.” Michael Lynch similarly writes: “To believe is just to take as true” (2004, p. 13), and Haack (1993, p. 192) writes that it is a truism about belief that “to believe that  $p$  is to accept  $p$  as true.” The list goes on.

cannot simply mean to *entertain the thought* that  $p$ . For one thing, we have a huge host of beliefs that we do not currently entertain, and perhaps even that we never have or never will entertain. For another, we might entertain the thought that  $p$  is true without believing that  $p$  (as when one *wonders* whether  $p$ ). Nor can we say that taking  $p$  to be true is to be *committed* to the truth of  $p$ . For one might expressly commit oneself to the truth of a proposition that one does not believe (e.g. in writing research proposals, one might find oneself tempted to make strong commitments to the truth of claims that one does not believe, as a way of convincing the grant committee that one's research is worth funding). And one can disbelieve propositions that one is committed to, as when one fails to see that one's explicit beliefs straightforwardly entail a proposition that one explicitly disbelieves.

Keeping these traps about taking  $p$  to be true in mind, we can say, at a minimum, that one takes as true all of those propositions that one explicitly and sincerely affirms. We can also say that one takes some things to be true, even if one does not currently have them in mind, and even if one has never consciously affirmed them. I take it to be true that Moscow resides in our solar system, although I have never (until now) ever consciously entertained that thought. Moreover, I took it to be true, even before I thought of it, although I have never (as far as I can recall) entertained that proposition in my mind. These reflections obviously do not settle all of the questions about the nature of belief, but they look plausible as far as they go.

The definition of belief as taking a proposition to be true connects up with the epistemic goal, because on a goal-based account of reasons, one has a reason to do or believe  $X$  only when and because doing or believing  $X$  promotes the achievement of some goal. So if one has an instrumental reason to believe that  $p$ , then there must be a goal that believing that  $p$  promotes, which explains why there is a reason to believe that  $p$ . The fact that believing that  $p$  is to take  $p$  to be true fits very naturally with taking truth to be the focus of the epistemic goal, on the instrumental account of epistemic rationality, because if truth-monism is correct, then the epistemic goal can help explain ordinary reasons for beliefs. A reason to believe that  $p$  is necessarily a reason to take  $p$  to be true; if reasons are goal-based, then there must be a goal that explains why there is a reason to take  $p$  to be true; so taking truth to be included in the epistemic goal allows the instrumentalist to explain how one can have a reason to believe  $p$  (by holding that achieving the goal of having true beliefs and avoiding false beliefs is part of the epistemic goal). Furthermore, if we do not include anything else in the epistemic goal, then *only* what bears on the truth and falsity of propositions will be able to play the role of an epistemic reason, because only what promotes the achievement of true beliefs and the avoidance of false beliefs will be an epistemic reason (i.e. will bear on the achievement of the epistemic goal) – and it is a very natural thought that whatever does not bear on the truth or falsity of a proposition is not an epistemic reason for belief. So we have some reason to include truth-achievement and error-avoidance – and nothing else – in the set of goods to be included in the epistemic goal.

All of that assumes that believing that  $p$  really is to take  $p$  to be true. There are two possible objections one might raise against this definition. For one thing, one might hold that it is not necessary for believing that  $p$ , that one take  $p$  to be true. Small children and animals can have beliefs, and (one might claim) they do not have the concept of truth.

It is not the purpose of this dissertation to weigh in on the debate on the nature of truth. However, it seems to me that correspondence theorists as well as deflationists at least ought to be able to agree that if one knows what it is like for the world to be one way rather than another – if one can *think* that the world is one way rather than another – then one has at least a minimal concept of truth. If that is correct, then any being capable of having beliefs will necessarily have a concept of truth, for even the simplest of beliefs represents some part of reality as being one way rather than another. And it then follows that, if small children and animals really do have beliefs, then they do have a concept of truth.<sup>38</sup>

The other objection that might be raised against this picture of belief is the pragmatist line that genuine beliefs have practical consequences; they are taken up in the active life of an agent. It is therefore not sufficient for having a belief that one take a proposition to be true; one's active life is the criterion of what one believes.

---

<sup>38</sup> On the other hand, perhaps small children and animals do not have full-fledged beliefs. I find Millar's view (2009, pp. 147-148) very tempting: we ought first to get clear on the clear cases of beliefs, which are the beliefs of reflectively aware adults, and then go on to determine whether others can have beliefs in that sense too, rather than to assume that small children and animals have beliefs in the same sense as adult humans do and start theorizing from there.

There are two points to make by way of response to this line of argument. First, it is not particularly plausible: one can have a propositional attitude without having it directly influence one's active life. It only does so if the proposition in question bears on things one is interested in acting on. Furthermore, the action that results from a belief in a given situation will depend on the agent's desires/values and her other beliefs. There is no simple, straightforward connection between a belief and an action. (Should I take my umbrella? Well, that depends on whether it is going to rain. And on whether I am aware of that. And on whether I know that umbrellas keep people relatively dry in the rain. And on whether I want to stay dry. And so on.)

The second point to make is that even if the pragmatist account of belief is correct, it does not follow that believing that  $p$  is not identical with taking  $p$  to be true. What follows is, rather, that there is a practical constraint on taking  $p$  to be true: in order to take  $p$  to be true, one must take it up in one's practical life.

To recap: to believe that  $p$  is to take  $p$  to be true. On the instrumental account, reasons are goal-based, and so a reason to believe that  $p$  must be goal-based, i.e. it must involve reference to a goal that believing  $p$  promotes. So a reason to believe that  $p$  must make reference to a goal that taking  $p$  to be true promotes. One such goal is obviously the goal of believing truths. The truth-goal therefore explains the possibility of reasons for belief, on the instrumental account, and so it makes good sense to include truth in the epistemic goal.

I would like to emphasize, however, that as I am not an instrumentalist, I do not want to seriously commit myself to any of the above views, except the account of belief as taking-to-be-true. I am only trying to give the best formulation of the epistemic goal, in order to paint the instrumental conception of epistemic rationality in its most favorable light. By and large, epistemologists assume that truth is valuable in some sense, and they assume that that is enough to ground a distinctively epistemic kind of value; the question is whether anything else is epistemically valuable too. (Sometimes the value of truth is argued for, e.g. by arguing that true beliefs tend to have better consequences (Foley 1993), or by arguing that valuing truth is partly constitutive of happiness (Lynch 2004), but arguments for the value of truth all by itself are hard to come by.) I have tried here to offer a reason for taking truth to be epistemically valuable (for an instrumentalist, at least) in the first place. If that case is successful, well and good. If not, then I am content to simply assume for now, along with everyone else, that truth is to be included in the epistemic goal.

One might, however, want to hold that coherence, simplicity (in some form or other), etc., ought also to be considered as evidence for propositions, and they ought therefore to be included in the set of epistemic goals. These doubts bring us to the third avenue of support for truth-monism: other epistemic values can be reduced to the value of truth. The main reason that coherence is so plausibly taken to bear on the epistemic status of a belief is that it *does* often bear on the truth or falsity of propositions. To recognize the evidential force that a high degree of coherence with a well-established explanatory

system confers on a proposition is perfectly consistent with truth-monism; the epistemic goal of achieving true beliefs and avoiding errors in fact *requires* us to take into account whatever considerations bear on the truth or falsity of our beliefs. If coherence has such a bearing, then truth-monism requires that we take it into account.

It is an open question whether it is possible to perform a similar reduction, going in the other direction: to reduce the value of truth to the value of coherence, simplicity, or some other such thing. (I am not aware of anyone who attempts such a reduction, except perhaps Rorty (1998, ch.1) after a fashion, but someone might.) It follows that the possible reduction of those values to the value of truth is not a knock-down argument for truth-monism. However, there is near-universal agreement that if anything is epistemically valuable, truth is, while there is no such agreement about, say, simplicity, so if other candidate epistemic values can be reduced to the value of truth, that is enough for my purpose here to take truth to be fundamental and those others as secondary.

So the case for truth-monism, in a nutshell, is: (1) some epistemic values cannot be epistemic goals, on pain of circularity; (2) the nature of belief, as taking a proposition to be true, entails that a reason to believe that  $p$  is a reason to take  $p$  to be true, so the epistemic goal will limit epistemic reasons to those that bear on a proposition's truth (which seems to be a good restriction to be able to have on possible reasons for belief – certainly most contemporary analytic philosophers agree with it); and (3) the epistemic value of other plausible candidate epistemic goals like coherence and simplicity can be reduced to the epistemic value of truth.

All of that is by way of identifying the most plausible elements of the epistemic goal. Bear in mind that I am not myself an instrumentalist, and so the formulation of the epistemic goal is only a means to an end: I am attempting to construct the most plausible instrumentalist position in order to show that epistemic instrumentalism even in its strongest form is not tenable. Still, I find the case for truth-monism about the epistemic goal to be fairly convincing.

### **4.3. Matters of Interest and Importance**

Recall Alston's revised formulation of the epistemic goal, as "maximizing true beliefs and minimizing false beliefs about matters of interest and importance" (2005, p.32). The reason instrumentalists will want to include the clause about interest and importance is to avoid the frequently-raised objection (e.g. by Sosa (2002), Engel (2002, ch.5), and Grimm (2009)) that there are true beliefs that seem to have no value whatsoever, and which most people could not possibly care about having. For example, a true belief about the number of grains of sand on all the beaches of the Caribbean is a true belief not worth caring about. Therefore, if the epistemic goal is to be a goal that matters for agents to achieve, it must be restricted to matters of interest or importance to people. Alston includes the interest and importance clause to avoid objections like these. Similarly, Piller holds that the existence of uninteresting truths is not a serious objection to the view that the truth is interesting, because the way that we should understand the claim *I am interested in the truth* is not as an interest "in any truth, obviously, but in the correct answers to questions that are of some concern to me" (2009a, p.195).

It is worth pausing a moment to reflect on just what it means for a truth to be interesting and/or important. Why do some truths count as interesting or important while others do not? Are there even any uninteresting or unimportant truths? Can being paid to discover the truth about some question make an otherwise uninteresting truth interesting for me?

It might help us here to consider the purpose for which the interest and importance clause was called in. We want to include interest and importance in the formulation of the epistemic goal, because there are some truths that just do not appear to be the least bit valuable to have. If there is a truth-goal that is of interest to people, it must not entail that people are interested in having *any* old true belief, because there are examples of true beliefs that appear to be entirely uninteresting. Because of such beliefs, if caring about the truth entails caring about any true belief at all, then very few people (if any) care about the truth in this sense.

I do not see the terms “interest” and “importance” defined anywhere by those who use them, but they have a fairly intuitive meaning. We can say that a truth is interesting to a person if it offers some satisfaction to the person’s curiosity. This is typically the case when you just want to know the truth about something. Perhaps you’ve begun counting the bricks in a walkway because you were bored, and now having begun, you want to finish the task and discover just how many there are. Someone might ask: why do you care? An appropriate response would be: no reason – I just want to know. Important

truths, on the other hand, would be those that have some sort of practical payoff, or are valuable in themselves.

Understood in this way, interest and importance can come apart. A true proposition can be important without being interesting. Imagine that a wealthy but eccentric philanthropist pays me a great deal of money to find out whether people who put on their left pant leg first also put on their left sock first. I am not the least bit curious about whether that is the case, but it can be important to me to find it out, because there is a practical payoff to doing so.

Similarly, a true proposition might be interesting to someone, when it satisfies some idle intellectual curiosity, without being the least bit important. For example, after having read too much epistemology, I once found myself wondering, just what is the 10<sup>th</sup> entry in the Wichita, Kansas telephone directory, anyway?<sup>39</sup> And yet, I think that whatever the entry is, it is entirely unimportant. I *just wanted* to know (at least a tiny bit – I never did look it up, so it must not have been a very great curiosity).

However, I imagine that most of the time, interest and importance go together. When someone is curious about whether some proposition is true, it becomes important to her to find it out. And when one recognizes that the truth about some matter is important, one often becomes interested in the truth of the matter.

A further point about interest and importance is that they both appear to be person-relative, at least in many cases. What one person finds interesting, another might

---

<sup>39</sup> Goldman's (1999, p. 88) example is about the 323<sup>rd</sup> entry, but I found myself wondering about the 10<sup>th</sup>. I still was not interested in the 323<sup>rd</sup>.

find utterly boring. (I have some interest in the paradoxes of material implication. I doubt that this is a common interest, perhaps even among philosophers.) And what is important to one person might be entirely unimportant to another. (For example, it is important to me whether I complete my dissertation: I care about it for its own sake, and also as a means to get a job in academia. But to someone who is cold and starving, I doubt that my dissertation matters at all.) Now, I do not want to rule out at this point the possibility that some truths are important in themselves, and so in a sense are not relative to persons, but it does seem that the importance of many if not most important truths is person-relative. I doubt that *all* true propositions are important in themselves (though we will see some arguments that they are in Chapter 7), but perhaps at least some are.<sup>40</sup>

A related point is that there could very well be some people out there who are interested in the truth, in the sense that for any old proposition that is true, they'd prefer to believe it than either to have a false belief about it or to have no belief about it at all. To such a person, the interest in truth would be boundless. But such a person is not the typical case; for most people (or at the very least, for me!), there are at least some truths that are not the least bit interesting.

Coming back to how we should understand the interest and importance clause: I think we can safely say that to add the interest and importance clause to the epistemic goal is to restrict the goal to those truths that either satisfy someone's intellectual curiosity or else have some practical importance for someone or are valuable in

---

<sup>40</sup> Perhaps all truths are important because true propositions say something about the world in way that false propositions do not. Perhaps they are all important because being concerned for *all* truths is an important element of happiness. We'll come back to arguments like these in Chapter 7.

themselves. Including this restriction makes some sense, since it would certainly be odd to evaluate people's beliefs in the light of an epistemic goal that does not satisfy some curiosity, or have some practical importance, or have some value in itself.

The problem with including the “interest and importance” clause in the formulation of the goal, however, is that it opens up the instrumental account to a serious objection: if one acquires good evidence for the truth of a proposition, and one therefore comes to believe that proposition, that belief can be epistemically rational, *even if* one does not care at all about the truth of that proposition.<sup>41</sup> If the National Sand-Grain Counting Center declares that, after years of careful, peer-reviewed study, the total number of grains of sand on all the beaches of the Caribbean is certain to be seven trillion, then my belief that that is the number of grains of sand on those beaches can be epistemically rational. However, if the epistemic goal includes the “interest and importance” clause, then the instrumental conception cannot accommodate the epistemic rationality of that belief, because the belief would not serve to achieve the epistemic goal.

The issue of apparently worthless true beliefs will come back in Chapter 7, as one of the main lines of criticism of the instrumental conception. The most obvious line of

---

<sup>41</sup> Ok, but is this a serious problem for a pluralist account of rationality, like Foley's, or of epistemic desiderata, like Alston's? Yes. If, as I argued in Chapter 1, the only interesting kinds of instrumental rationality are the objective and the subjective varieties as I have explained them, then there is no interesting kind of instrumental rationality that can accommodate the kind of case at hand here. It follows, then, that allowing that cases like these are epistemically rational goes beyond what Foley and Alston want to allow as possible types of rationality (Foley) and epistemic desiderata (Alston), because the different varieties that they want to allow are still all instrumental in nature.

Foley's pluralist view of epistemic goals, by the way, is consistent with his monistic formulation of the truth-goal, which I endorsed as the most plausible formulation of the epistemic goal in Chapter 4. Foley takes it that there are different senses of rationality that correspond to different goals. But his preferred type of rationality is that characterized by the goal of now achieving true beliefs and now avoiding false beliefs.

response open to the instrumentalist is to hold that all truths really are worth caring about (really are either interesting or important), at least a little bit (cf. Foley 1993, p.17; Lynch 2004). We will examine that response more closely later. For now, what is important is that, according to the subjective instrumental conception of epistemic rationality, it is the epistemic goals that agents care about achieving that determine the epistemic rationality of their beliefs. So according to the subjective conception, it is either truths that people find interesting or that people think are important that matter for the purpose of determining the epistemic rationality of their beliefs. And, according to the objective instrumental conception, it is the epistemic goals that are good to achieve independently of what agents care about that matter for the determination of the epistemic rationality of people's beliefs. So, according to the objective instrumental conception, it is truths that are in fact important that matter for the purpose of determining the epistemic rationality of people's beliefs.

What this means is that adding the interest and importance clause to the formulation of the epistemic goal is just to make explicit what is already implicit in talking in terms of the subjective and the objective instrumental conceptions of epistemic rationality. It is not just any truth-related goal that matters for the determination of epistemic rationality; it is only those goals that agents in fact care about achieving, or else that are good to achieve whether or not agents care about them, that matter. That is just another way of saying that it is only interesting and/or important truths that matter. So I

do not think that adding the interest and importance clause to the epistemic goal makes any difference.

It is important to keep in mind, though, that epistemic evaluations can be made about any beliefs whatsoever, even extremely trivial or apparently unimportant ones. So the epistemic goal must include the achievement of truths and avoidance of errors even in such cases, if it is to provide us with a complete account of epistemic rationality.

Does it beg the question to say, as in my imaginary case of the National Sand-Grain Counting Center, that beliefs that are held on the basis of good evidence are epistemically rational? No. Instrumentalists typically do not want to deny the importance of evidence in an account of epistemic rationality. Foley (1993, ch.1) emphasizes the close connection of evidence and the epistemic goal; on his account, the epistemic goal must not come apart from evidence – at least, not in most cases. Indeed, this is the reason for Foley’s synchronic restriction on the epistemic goal:

For example, suppose a proposition P involves a more favorable assessment of my intellectual talents than the evidence warrants, but suppose also that believing P would make me more intellectually confident than I would be otherwise, which would make me a more dedicated inquirer, which in turn would enhance my long-term prospects of having an accurate and comprehensive belief system. Despite these long-term benefits, there is an important sense of rational belief, indeed the very sense that traditionally has been of the most interest to epistemologists, in which it is not rational for me to believe P. (2005, p.317)

Here, Foley uses evidential standards to argue for his preferred formulation of the epistemic goal.

Foley does allow that evidence and epistemic reasons for belief can come apart in exceptional cases: there can be cases where there is evidence for the truth of a

proposition, but no epistemic reason to form the belief in question. Such cases are those where the formation of a belief can itself simultaneously undermine the evidence for the belief. Here is an example of such a case:

### **A Shocking Thought-Experiment**

We are doing a science experiment. I hook you up to a brain scanner, so that I can know whether you believe that I am going to shock you. I tell you that if you form the belief that I am going to shock you, then I will not shock you; if you believe that I will not shock you, then I will shock you; and if you withhold judgment, I will shock you. You begin with no belief about whether I will shock you. But then you see that, in that state of non-belief, you have evidence for thinking that I will shock you. But if you proceed to form the belief warranted by that evidence, the evidence itself disappears: you will then believe that I am going to shock you, and when you have that belief, I will not shock you.

This is an epistemically unlucky case. There is evidence to be had, but when a belief is formed in accord with the evidence, the evidence changes. In such cases, Foley thinks, there is no epistemic reason to form a belief, despite the available evidence, because no matter what belief you form, it will fare badly in light of the epistemic goal.

This is an exceptional case, though, and even in this case, we should bear in mind that the reason that the available evidence fails to generate a reason for belief is that there is also evidence to the effect that the proposed belief cannot coexist with the evidence that

is currently available for it. To accommodate cases like this one, Foley holds that evidence and the epistemic goal are *subjunctively* linked:

Having sufficient evidence for a proposition gives you adequate epistemic reason to believe it, unless believing the proposition would itself undermine the evidence. Correspondingly, if you don't have sufficient evidence for a proposition, then you don't have an adequate epistemic reason to believe it, unless believing it would itself create adequate evidence for the proposition. (1993, p.30)

Other instrumentalists also want to emphasize the importance of evidence. Alston (2005, p.92) writes: "I take it that a ground for a belief could not be called 'adequate' in any natural sense unless it does have some bearing on the truth of the belief." Leaving open the question of whether there could be adequate non-evidential grounds for belief (Alston would think so), we can note that evidence does bear on the truth of beliefs, so that good evidential grounds are adequate grounds, on Alston's account. And if there is good evidence against a belief, then to hold the belief anyway is epistemically irrational.<sup>42</sup> Similarly, Vahid (2003, p.85) writes that "Unlike truth, justification is perspective [sic] i.e., it is determined relative to the cognizer's evidence." Appealing to cases where beliefs are held either on the basis of good evidence, or else against good evidence, as test cases to see whether the instrumental conception can handle them, does not beg the question because instrumentalists themselves think that evidence is very important in an account of epistemic rationality.

---

<sup>42</sup> Alston would not put the point this way, since he does not talk in terms of rationality or justification. He might instead say that there are good grounds for rejecting the belief.

#### 4.4. Synchronic vs. Diachronic

What we have so far is that the epistemic goal involves achieving the truth and avoiding error, and that it applies to all propositions. The question at hand now is whether the epistemic goal is to be understood synchronically or diachronically – that is, whether the goal is *now* to have true beliefs and *now* to avoid false beliefs, or whether the goal is to have true beliefs and to avoid false beliefs over an extended period of time.

I am inclined to side with defenders of the synchronic formulation. The main reason for adopting the synchronic conception is that the future epistemic consequences of one's beliefs shouldn't have a bearing on the *current* epistemic status of those beliefs. If the epistemic goal is diachronic, so that what makes a belief epistemically rational is whether it promotes (/tends to promote/would be taken on reflection to promote) the goal of having, say, a comprehensive body of beliefs with a favourable truth-falsity ratio over the long haul, then if holding a particular belief now against the available evidence would have good epistemic consequences in the future, that belief would count as epistemically rational. But holding beliefs against the available evidence is paradigmatic of epistemic irrationality. So the epistemic goal must be synchronic.<sup>43</sup>

This argument is a *reductio* against the diachronic conception, which works by showing that the diachronic formulation of the epistemic goal entails that certain beliefs that should be counted as epistemically irrational are epistemically rational (and vice versa). A defender of the diachronic conception might want to respond by offering a

---

<sup>43</sup> Cf. Foley (1987) and David (2001), who make similar arguments.

similar *reductio* going the other way: we know that certain beliefs are epistemically rational, because they fare well in light of the diachronic goal, but the synchronic goal counts them as epistemically irrational. This response on behalf of the diachronic formulation of the epistemic goal is not a serious problem, though, because the argument for the synchronic restriction does not work by assuming that the problematic beliefs in question are epistemically rational due simply to the fact that they fare well in light of the synchronic goal (or irrational due to the fact that they fare poorly in light of the synchronic goal).<sup>44</sup> It works by assuming that the beliefs in question are epistemically rational because they are based on good evidence (or irrational because they contradict good evidence). Then the argument proceeds by noting that the synchronic formulation of the epistemic goal does a better job of categorizing those beliefs as epistemically rational or irrational than the diachronic goal. So the argument that is run against the diachronic formulation cannot be turned around and used against the synchronic formulation.

There are three other lines of objection that might be brought to bear against the synchronic formulation of the epistemic goal, though. First, if true beliefs as such are epistemically valuable, then future true beliefs ought to be counted as part of the epistemic goal. The response to this objection is easy enough. We already have a

---

<sup>44</sup> Or at least, the argument *need* not work in that way. One of Foley's arguments against the diachronic goal does seem to work this way, because he claims that future epistemic benefits and harms of current beliefs need not reflect the evidential status of those current beliefs, while synchronic or "purely" epistemic goals go hand in hand with evidence – and he also claims that synchronic epistemic goals impose a restriction on accounts of evidence (1993, p. 20). So this argument of Foley's for the synchronic formulation of the epistemic goal is circular. The way to escape the circularity is to make the account of evidence independent of the synchronic epistemic goal; then the appeal to evidence allows us to prefer one formulation of the epistemic goal over another. Foley's (2005, p. 317) argument for the synchronic formulation of the epistemic goal, cited above (section 4.3), proceeds this way, and avoids circularity.

distinction in place between epistemic goals and epistemic values, and as we have just seen, we have a good reason to keep the value of future true beliefs firmly outside of the epistemic goal: including them gives rise to cases of obvious epistemic irrationality that the instrumental conception would have to count as epistemically rational. So what we can do is count them in the category of epistemic values, but not in the epistemic goal.

The second and third lines of objection to the synchronic formulation that I have in mind come from Vahid (2003). He argues, for one thing, that epistemic justification supervenes on proper belief-forming processes (or on adequate grounding – Vahid is not particular about the type of causal history that counts, only that it is a type of causal history that counts), and proper belief-forming processes or groundings for beliefs are those that promote the maximization of true beliefs and the minimization of false beliefs over the long run. Therefore, according to Vahid’s argument, beliefs that themselves promote the long-term achievement of truths and avoidance of falsehoods, but which are formed by an improper belief-forming process (e.g. one that allows belief-formation in the face of contrary evidence) do not count as justified. It follows that the argument that I recounted above for a synchronic formulation of the goal – that including the achievement of future true beliefs in the epistemic goal will get the extension of epistemically rational beliefs wrong – is incorrect, because the beliefs that are alleged to promote the achievement of the diachronic goal even in the face of contrary evidence do not really promote the diachronic goal after all.

Vahid also points out that adopting a diachronic formulation of the epistemic goal gets around an objection that Maitzen (1995) raises to making the epistemic goal a matter of having true beliefs and avoiding false beliefs. Maitzen argues that given the truth-directed epistemic goal, it is hard to see how there is any room for justified false beliefs, or unjustified true beliefs. A true belief always promotes the goal, while a false one always hinders it. Vahid argues that Maitzen's argument only goes through if the epistemic goal is taken synchronically. If the goal is diachronic, false beliefs can be epistemically justified, if they are produced by processes that have (or if they are grounded on evidence that has) positive overall epistemic consequences, because, again, Vahid holds that justification comes out of the diachronic truth-conduciveness of the causal histories of beliefs, and diachronically truth-conducive types of causal histories of beliefs can generate false beliefs in particular unlucky cases.

There are convincing replies to be made on behalf of the synchronic view of epistemic goals. First of all, Vahid's response to the argument recounted above in support of a synchronic epistemic goal does not work. His argument rests on the assumption that a belief that is formed in an improper way (for instance, when a belief is formed contrary to the available evidence, or by an unreliable process) does not promote the diachronic epistemic goal, because that way of forming beliefs does not promote the achievement of truths and the avoidance of errors over the long run. However, if the epistemic goal is to achieve truths and avoid falsehoods over the long run, then presumably an agent who considers the long-term epistemic benefits and harms that holding a belief  $p$  will have,

who recognizes that believing  $p$  will generate many true beliefs and very few false ones over the long run, and who therefore forms the belief that  $p$ , has done an excellent job in forming the belief, vis-à-vis the diachronic epistemic goal. We would have to count  $p$  as adequately grounded, *even if* the currently available evidence indicates that  $p$  itself is false.

Vahid's move, then, of shifting the focus in an epistemic evaluation from how well target beliefs *themselves* promote the long-term epistemic goal to how well the processes that *generate* the beliefs promote the long-term epistemic goal, is open to objection. There is at least one type of belief-forming process that is perfectly good, considered in light of the diachronic epistemic goal, but which in fact generates clear cases of epistemically irrational beliefs. That process is the one that moves from a careful consideration of future true beliefs and false beliefs that will follow from believing a target proposition  $p$ , to the formation of the belief that  $p$ , even if the currently available evidence tells against the truth of  $p$ . That process can generate beliefs which the available evidence indicates are false, but which are causally necessary for achieving many true beliefs or avoiding many errors later. Such beliefs are not epistemically rational, since they are formed contrary to the available evidence, but they are formed by a process that promotes the diachronic epistemic goal.

To put the objection here in other words, Vahid assumes that belief-forming processes that allow the formation of beliefs contrary to the available evidence are not

diachronically truth-conducive. The only argument Vahid offers in support of this claim is the following:

Since justification is perspectival, the available evidence  $e$  at  $t$ , being incomplete, might indicate that our belief  $p$  is more likely to be false, even if, as it turns out, it is in fact true. Relative to  $e$ , the belief  $p$  is not adequately grounded. So believing  $p$  is not a good thing from the epistemic point of view, and thus, unjustified. However, we have to treat like cases alike. This means that in the future we have to treat other inadequately grounded beliefs as unjustified. As a norm, however, most of the inadequately grounded beliefs turn out to be false. (2003, p.86)

The idea is that most beliefs that are formed against the evidence (which is a way of being inadequately grounded) are false, so forming beliefs against the evidence is not diachronically truth-conducive.

No doubt Vahid is correct, in the ordinary case: because evidence bears on the truth of a proposition, holding beliefs against the available evidence will generally not be truth-conducive over the long run. However, as I have tried to show, there is at least one belief-forming process that is diachronically truth-conducive, which allows the formation of belief in the face of contrary evidence. That is the process that begins with the conscious consideration of the long-term epistemic benefits and harms that will follow if one holds the belief, and if the future epistemic benefits outweigh the harms, then the process issues in the formation of the belief even in the face of evidence against the truth of the target belief itself. Because of the possibility of processes like this, I am arguing, believing in accord with the evidence is not necessarily tied to diachronic truth-conducive causal histories.

Regarding Vahid's other line of argument, that we need a diachronic epistemic goal in order to avoid Maitzen's objection that the truth-goal does not allow for the possibility of justified false beliefs and unjustified true beliefs: it is not clear that only a diachronic epistemic goal can avoid that result. A synchronic epistemic goal might be able to avoid the objection, by opting for an account of epistemic justification according to which a belief is epistemically justified for a subject if and only if she *thinks* that it promotes the epistemic goal, or she has good *reason* to think that the belief promotes the epistemic goal, or something along those lines. Maitzen's argument only goes through if it is those beliefs, and only those beliefs, that *in fact* promote the achievement of the epistemic goal (i.e. are true), that count as epistemically justified. If we do not hold epistemic justification to be about in fact achieving the epistemic goal, but rather about doing one's best to achieve the epistemic goal, doing what one *believes* will achieve the goal, doing what one *on reflection* will achieve the goal, etc., then justified false beliefs as well as unjustified true beliefs are possible. So we do not need to move to a diachronic epistemic goal to avoid Maitzen's objection.

#### **4.5. Summary: The Epistemic Goal**

The formulation of the epistemic goal that falls out of the foregoing discussion is essentially Foley's (1987). It involves truth-achievement and error-avoidance, it applies to all propositions, and it is synchronic. With this account of the epistemic goal in hand, we can now move on in the following chapters to consider some objections to the instrumental conception.

## **PART II: Objections to the Instrumental Conception**

## **Chapter 5: Proper Functions and Epistemic Rationality**

Let me begin by summing up what we have done so far. In the first chapter, I introduced the instrumental conception of epistemic rationality, as the view that the epistemic rationality of doxastic states depends on the content of some relevant epistemic goal(s). The two ways of filling in what makes an epistemic goal relevant for the evaluations of people's beliefs are subjective and objective. The subjective instrumental conception is that the goals that matter are those that agents care about achieving. The objective conception is that the epistemic goals are valuable irrespective of what subjects want to achieve. In the second chapter, we saw that a number of epistemologists espouse the instrumental conception, and also that there are a number of critics of that approach. In the third chapter, we went through a number of reasons for accepting the instrumental conception, and I argued that those reasons are not as strong as they might appear (although I did not argue that they provide no support at all). In the fourth chapter, we saw various ways to characterize the epistemic goal. The idea was to determine the most plausible formulation of the goal, in order to give the instrumental conception the best case that we could. I argued that the most plausible way to characterize the epistemic goal is essentially Foley's: now to believe truths and now to avoid believing falsehoods.

This chapter and the three which follow it are concerned with setting out the objections to the instrumental conception. As we have seen, there are two broad types of instrumental conception of epistemic rationality to consider. The subjective conception holds that the epistemic rationality of beliefs depends on the epistemic goals that

epistemic agents care about. The objective conception holds that epistemic rationality depends on epistemic goals that are good to achieve, whether or not anybody cares about them. The present chapter is concerned with one way to be an objectivist about the epistemic goal: the *proper function* account of epistemic rationality. There are some epistemologists who want to ground epistemic rationality (or whatever their preferred terms of epistemic appraisal may be) in the proper function of either beliefs or cognitive systems. In what follows, I explain that way of attempting to analyze epistemic rationality, and then I go on to raise a number of objections to that view.

## **5.1. The Concept of a Function**

### **5.1.1. Accounting for rationality via functions**

We can begin by considering the possibility of grounding instrumental rationality in the proper function of either beliefs or cognitive systems. Some epistemologists are explicit about wanting to ground their analyses of their target epistemic concepts in the selected-effect notion of a proper function, a conception of function introduced into the literature primarily by Millikan (1984). (We will get to Millikan's concept of a function shortly.)

The idea is that the epistemic goal, now to believe truths and now to avoid errors, is what (at least part) of our cognitive systems have as their proper function. So that is why the epistemic goal is valuable to achieve.

I am not optimistic about the prospects for a proper-function account of rationality or justification. In the sub-section coming up, I will explain the concept of the proper function of an organ or trait. After that, I will set up the attempt to employ the notion of

proper function to ground epistemic rationality (or other epistemic terms of appraisal).

Finally, we will see some problems that arise for that way of doing epistemology.

### 5.1.2. What is a function?

In the literature on functions in the philosophy of biology, philosophy of mind, philosophy of language, and epistemology, there are two main function-concepts. First, there is the etiological, or selected-effects, concept of function. Sometimes such functions are called Millikan-style functions, after Ruth Millikan (1984; 1989), a prominent champion of this concept of function. Others who defend and make use of the selected-effects concept of function include Neander (1991a; 1991b), Godfrey-Smith (1994; 1996), and Perlman (2010). Details vary among the different accounts, but the basic idea is as follows:

**(Basic) Selected-Effects Function:** an item (organ or trait) X has function F just in case the ancestors of X did F, and the fact that they did F caused or explains the fact that Xs exist and do F now.

Selected-effects functions so characterized in fact cover more than merely those functions that have been naturally selected for,<sup>45</sup> but theorists typically appeal to it as explaining how natural selection gives rise to proper functions. The idea is that an organ or trait has a proper function if it was naturally selected for having the effect in question. The heart, for

---

<sup>45</sup> Strictly speaking, artifactual functions are also selected-effects functions, because artifacts are typically created precisely in order to produce some effect. Also, some types of effects that ancestor traits or organs had are quite obviously not proper functions, e.g. in cases in which the fact that the ancestors of X did F, and that fact provides an explanation of the fact that Xs exist and do F now, but the explanation involves some important causal deviance. But we can set that issue aside for now (causal deviance is a general problem for causal analyses anyway), and give the selected-effects analyses the benefit of the doubt.

example, has pumping blood as its proper function, because hearts in the past pumped blood, and the fact that they did so made the organisms in which they existed more fit for survival and reproduction than organisms without hearts that pumped blood (or with hearts that did not pump blood as well). The fact that hearts used to pump blood therefore explains the existence of hearts that pump blood in organisms today, and so pumping blood is the proper function of hearts.

The second function-concept in the literature follows Cummins (1975). Sometimes this kind of function is called a systematic function, but often it is simply called a Cummins-function. The idea behind Cummins-functions is the following:

**Cummins-function:** an item (organ or trait) X, that is part of a system Y, has function F just in case X does F and F contributes to the overall fitness of Y (or more generally, X's doing F contributes to some important effect E of Y's).

Lewens (2004) and Cummins and Roth (2010) have recently offered defenses of versions of Cummins-functions. An example of a Cummins-style analysis of a function is to say that the heart has the function of pumping blood, because the heart is a proper part of an organized system, and the heart's pumping blood contributes to the survival of that system.

### **5.1.3. The aim of belief and proper functions**

A number of epistemologists have become interested in the aim-of-belief thesis, the view that beliefs have a proper aim, and that their aim is importantly truth-related. However, the claim that beliefs aim at truth is very often taken to be a metaphor that needs to be

cashed out, because beliefs (not being the kinds of things that can *intend* to achieve an aim) cannot really *aim* at anything. Pascal Engel, for example, writes: “beliefs do not aim at anything by themselves, they do not contain little archers trying to hit the target of truth with their arrows” (2004, p.77). Engel holds that, at best, the aim of belief is a metaphor to be cashed out in terms of the aims of believers. Ralph Wedgwood independently employs the same image, though he interprets the metaphor of aiming differently:

It is often claimed that belief aims at truth. Indeed, this claim has often been thought to express an essential or constitutive feature of belief. But this claim is obviously not literally true. Beliefs are not little archers armed with little bows and arrows: they do not literally “aim” at anything. The claim must be interpreted as a metaphor.

I propose to interpret this claim as a *normative* claim – roughly, as the claim that a belief is correct if and only if the proposition believed is true. (2002, p.267, italics in original)

Others who hold that beliefs are not capable of literally aiming at truth, or at least that aim-of-belief-talk is metaphorical and in need of explanation, include Millar (2009), Gibbard (2007), Owens (2003, p.289), Graham (2011a, p.56), Dreier (2010, p.159), Cruz and Pollock (2004, p.137), Lynch (2004, p.13), and Vahid (2006). Different ways of cashing out the metaphor include taking it to stand for an aim of *believers* when they hold beliefs (our subjective instrumental conception), or to mean that believers *ought* to aim at the truth when they hold beliefs (our objective instrumental conception), or that most justified beliefs really *are* true, or that truth is a good property for beliefs to have, or that the believer is *sensitive* to the truth of the belief.

However, even if aim-talk is best taken to be metaphorical, as applied to beliefs and cognitive systems, there are a number of epistemologists who make use of the related

notion of a proper function to ground their analyses of their preferred epistemic concepts. Lycan (1988, p.144), Bergmann (2006b, ch.3), Burge (2003), Plantinga (1993) and Graham (2011a, 2011b) all employ the notion of a proper function in their epistemologies. (Although Graham is clear about taking aim-of-belief-talk to be metaphorical, because beliefs do not have intentions, he takes the analysis in terms of proper functions of cognitive systems to be the way to cash out the aim metaphor, since biological proper functions do not rely on intentions.)

In what follows, the concept of function that we will mainly be working with is the selected-effects concept. For one thing, that is the dominant concept in the literature on functions. For another, Cummins-functions, interesting though they may be, are not the kind of function that that epistemologists tend to employ (cf. Graham 2011b), since they are quite explicitly intended to be non-normative. Recall that a Cummins-function is just the actual effect produced by an item that exists in a system, where the item's effect contributes to the production of an important effect (or towards the survival, which is one sort of important effect) of the system in question. When an item fails to have the effect in question, that item no longer has that function. A heart that stops pumping blood no longer has the Cummins-function of pumping blood. A sperm that does not fertilize an egg simply does not have the Cummins-function of fertilizing an egg. As we will see shortly, one of the key aspects of selected-effect functions is that they can account for malfunctions, but the Cummins-style of functional analysis avoids talking about malfunctions and normativity in nature. Indeed, Cummins and Roth (2010) take that to be

a virtue of this style of functional analysis; they caution the selected-effect functionalist against reading normativity into the functions of organs in nature, since that way of thinking slides into pre-Darwinian vitalism or supernaturalism about the designer of organs and biological systems. (Plantinga (1993), too, takes function and design in nature to require a supernatural designer, but rather than taking that to be a mark against employing the notion of proper function and a design plan, he takes it to be a mark in favour of metaphysical supernaturalism.) Perhaps we might try to revise Cummins's analysis of functions, so that they can account for malfunction (given that contributing to an important effect or to overall fitness might seem somehow to be normative). But (1) one of the aims of employing Cummins-functions is to avoid talk of normativity in nature. Bad things can be fit for survival, so fitness and selected effects are not always good. Fitness aside, it is not clear how the important effects of a system could even be distinguished from its unimportant effects, without making reference to the interests of agents. And (2) people who employ proper-function-talk in epistemology tend to have the selected-effect concept of function in mind, so that is the one we will focus on.

## **5.2. Proper Functions and Cognitive Systems**

There are a number of problems facing the proper-function analysis of epistemic rationality. One class of such problems, which will be the substance of sections 5.2.1 to 5.2.3, has to do with the fact that cognitive systems need not in fact have proper functions in the relevant sense, and if they do not in fact have functions, then functions do not ground epistemic rationality. In other words, intuitively, beliefs can be epistemically

rational or irrational, even for agents whose cognitive systems do not have functions in the relevant sense.

The second class of problems has to do with the fact that proper functions do not ground any serious kind of normativity. Sections 5.3.1 to 5.3.3 will set out these objections. The basic idea is that, although selected-effects accounts of function do allow for talk of a kind of malfunction, that is not real normative talk (in a sense to be explained below). If that is correct, then proper functions do not ground an analysis of epistemic rationality, which requires that epistemic reasons are normative for agents. People are rationally at fault if they do not form beliefs in accord with the epistemic reasons available to them, but there is no general rational requirement that people avoid doing things that conflict with the proper functions of their organs or traits.

### **5.2.1. Objection 1: natural selection, populations, competition**

This section will develop the first of three arguments intended to show that grounding epistemic rationality on the proper function of cognitive systems does not work. This first objection to proper-function analyses of epistemic concepts is that natural selection can only operate on whole populations, not on individuals, and it can also only operate where there is variation among the members of a population with respect to some heritable trait (Lewens 2004, ch.2). Consider the following example. Suppose that there is some population of birds with an average wingspan of one foot. These birds are the natural prey of a larger, stronger bird. These birds also eat a certain kind of worm that is plentiful in their environment. One day, the worms undergo a mutation which causes the birds who

eat them to have offspring with a smaller average wingspan. All of the birds proceed to eat these worms, and the next generation of the birds has a smaller average wingspan. As a result of this reduced wingspan, the birds are now more manoeuvrable, and better able to avoid their predators.

In this case, we have a trait in a population that confers survival value on its possessors, namely a smaller wingspan, where the trait was not selected for, because there was no variation among the members of the bird species with respect to the trait in question. So this is a trait with survival value but without function, because function requires that a property be selected for, not merely that the trait confer survival value. What this shows is that simply possessing survival value is not enough to ensure that a trait has been selected for.

The objection that results from this characteristic of natural selection is that it could be the case that our cognitive systems were not naturally selected for against competitor cognitive systems; cognitive systems could have all evolved simultaneously (not to say instantly, just that perhaps cognitive systems all mutated, grew, and learned together in the same way), without competition. Or it could be that certain brain structures evolved for reasons entirely separate from the production of true beliefs, which merely had the side effect of producing true beliefs. When we look back and see that our cognitive structures *did* evolve in such a way as to be more or less reliable in the production of true beliefs, it seems natural to infer that our cognitive systems evolved the way they did *because* they produce true beliefs, given that true beliefs are so useful to us

(and nice to have besides). But bear in mind Gould and Lewontin's admonition: "The immediate utility of an organic structure often says nothing at all about the reason for its being" (1979, p.593). Organic structures can come into being, and they can be selected for, for some reason that has nothing whatsoever to do with why that structure is valuable to us *now*, even if what it does now is extremely useful. It simply does not follow that an organic structure is selected for some effect E, from the mere fact that its producing E now is useful, or from our imagining it to have been useful in the past.

A related point is that some traits or structures are not selected for the effects that they themselves have, but catch a free ride by being attached to other traits or structures that are selected for. One way in which the evolutionary history of our cognitive systems could undermine the proper-function account of epistemic rationality is therefore the possibility that our truth-conducive cognitive systems were not selected for, but are simply catching a free ride on some other traits or systems that have been selected for. This possibility would be true, for example, if the following combination of theses were true: (1) consciousness is an essential part of our cognitive systems and our beliefs; (2) consciousness is merely epiphenomenal; (3) no epiphenomenal trait can be selected for.

It is therefore possible that our cognitive systems do not have selected-effect functions. So, if our analysis of epistemic rationality rests on the proper function of our cognitive systems, then our analysis must be open to the possibility that none of our beliefs are epistemically rational or irrational. But our analysis should not be open to that possibility, since we have paradigm instances of beliefs that are epistemically rational –

the belief in the existence of one's hands, for example – that ought to count as rational whether or not our cognitive systems have the right kind of evolutionary history. (Or at least, we would need *very* good reason to think that none of our beliefs are epistemically rational or irrational, and speculative evolutionary history does not give us such a reason.)

The objection I am giving here is distinct from another related objection, which the proper functionalist can handle. This other objection is that beliefs can be produced in accord with the proper function of a cognitive system, and be epistemically irrational because they were so produced. Such beliefs are possible, if cognitive systems have proper functions like the pursuit of happiness or the avoidance of pain. Consider, for example, the fact that people tend to systematically overestimate their own prospects, as well as the way that they are thought of by others (see e.g. Taylor and Brown 1988; 1994). Selection pressures could very well have designed it so that we tend to have such beliefs, regardless of whether they are true. If that is the case, then beliefs can be produced in accord with the proper function of a cognitive system, but be epistemically irrational.

Plantinga offers an adequate response to this objection. The response is to distinguish parts of the cognitive system that have the production of true belief as a proper function from those other parts of the cognitive system that have other cognitive functions, such as happiness, survival, etc. It is only the proper function of those parts of the cognitive system that have the production of true belief as their proper function that determine the epistemic status of beliefs.

Plantinga's reply deals with this specific objection, but it does not deal with the more general objection that I am raising, that if it turns out that cognitive systems do not have the right kind of history, then they do not have proper functions at all. If there are no proper functions of our cognitive systems at all, then there are no segments of our cognitive systems that have the production of true belief as their proper functions. But let me be clear: I am not suggesting anything about how our cognitive systems have in fact evolved. I am only arguing that the possibility that they evolved in such a way as to undermine the attribution of proper functions to them ought to be enough to convince us that we cannot ground epistemic rationality in proper function.

### **5.2.2. Objection 2: cognitive systems, before they were selected for**

A second, related, objection to grounding epistemic rationality in proper functions is that even if our cognitive systems do have the right kind of evolutionary history to have selected-effect functions, that would not be the case for the first beings with cognitive systems like ours. Suppose that there was a time when there were some early humans with cognitive systems well suited to producing true beliefs, and some other early humans with cognitive systems ill-suited to producing true beliefs. These are the kinds of circumstances that could have given rise to our cognitive systems' being selected for, because there is a competitor over which our cognitive systems could have been selected. However, if the analysis of epistemic rationality in terms of functions is correct, then we would not be able to say that those early humans with cognitive systems like ours were ever epistemically rational or irrational in their beliefs, because their cognitive systems

(not yet having been selected for) lacked proper functions. But surely, it is because we have the cognitive systems that we do have, that our beliefs come up for evaluation as epistemically rational or irrational. The same ought to go for all beings with cognitive systems like ours, including our early ancestors, *before* our cognitive systems were selected for over the competition.

### **5.2.3. Objection 3: Swampman**

This section develops the third argument intended to show that epistemic rationality is not to be grounded on proper functions. This third objection is a version of the familiar Swampman objection to teleosemantics, the theory according to which the content of propositional attitudes is determined by their proper functions. The Swampman case, due originally to Davidson (1987), and employed by Sosa (1993) as an objection to Plantinga's proper functionalist epistemology, goes like this. Imagine that a bolt of lightning in a swamp had the freak effect of producing a being that is molecule-for-molecule identical to you. It has organs like yours, DNA like yours, conscious experience like yours. Surely, we should say, your doppelganger has beliefs and other contentful states. But it does not have the appropriate kind of history for having proper functions. So contentful states must not require proper functions.

The same objection can be run against the proper function account of epistemic rationality. If you have a molecule-for-molecule copy that just came randomly into existence, then your copy does not have the right kind of history for having proper functions. Its heart lacks the function of pumping blood; its hands, the function of

grasping; and its cognitive systems, the function of generating true beliefs. But surely, such a being could have epistemically rational and irrational beliefs. Indeed, since it is your doppelganger, it ought to have the very same beliefs as you, and it is not clear that we should deny its beliefs the status of being rational or irrational. Suppose I ask it, “why do you think that you have hands?” It looks at me in an annoyed fashion, then replies, “for the same reason you think you have hands: I can see and feel them, and I can use them to pick things up.” Now, surely, if my own belief that I have hands is rational, based on reasons like those, then I must be committed to saying that my doppelganger’s belief is also rational.

The Swampman argument has its champions and its critics. Graham (2011b) gives voice to two important objections. First, Swampman does not have a mind; he has no mental states with propositional content. Assuming a causal account of meaning and reference, Swampman might have the *feeling* of thinking about the very same things that his doppelganger is thinking about, but his thoughts are not *about* anything at all: “Swampman is merely a physical duplicate; he’s a counterfeit mind, a counterfeit human being. He’s a fake” (2011b, p.19, emphasis in original).

Second, even if Swampman has a mind, his beliefs do not enjoy epistemic entitlement (Graham’s preferred epistemic concept). Graham holds that there are two species of warrant, where “warrant” is understood as some positive epistemic status that attaches to beliefs and to changes of beliefs. The first species is *reason*; the second, *entitlement*. On Graham’s account, roughly, reasons and justification go hand in hand,

and they are internalist: you can only justify your beliefs with the use of reasons that are at your disposal (ibid. pp.4, 6). Entitlement, on the other hand, is any warrant that is not a reason; it is externalist, and is a matter of the normal functioning of cognitive systems or belief-forming processes that have the function of reliably forming true beliefs (ibid. p.15).

In order to deny that Swampman's beliefs enjoy entitlement, Graham highlights the distinction between justification and entitlement. Swampman can have justification for his beliefs, since he can cite reasons (the very same reasons that would be cited by his doppelganger) in support of them. So, "He can justify his beliefs. Using 'entitlement' that way, Swampman's beliefs come out entitled. But that's not my use" (ibid. p.19). Graham holds that what we need to do in order to get a counterexample to *his* account of entitlement is to construct a case of a philosophical Swamp-zombie: a randomly generated near-duplicate of a human, with perceptual belief-forming processes mostly like ours, but who has no justifications (no reasons it could cite in support of its beliefs), and no rich phenomenal experience. The Swamp-zombie is a freak cosmological accident. Its cognitive systems have no history, so no etiological functions. It does not enjoy entitlement for its beliefs.<sup>46</sup>

A successful use of the Swampman objection, now, has to be able to show both that Swampman has a mind, with real beliefs and other propositional attitudes, and also to

---

<sup>46</sup> Graham switches from explicit arguments to rhetorical questions about the Swampman case at this point. He does not explicitly claim that the Swamp-zombie does not enjoy entitlements, but I take that to be his point.

show that Swampman's beliefs enjoy the relevant kind of epistemic properties. To those tasks I now turn.

First of all, Swampman is a molecule-for-molecule replica of a real human. He is the very same, right down to the DNA. What is even more important, he has phenomenal content identical to a normal human, and it certainly seems to him that he has beliefs about objects in the world. We ought to admit that he has a real mind, not a mere counterfeit.

If we accept the causal theory of reference, though, we have to admit that Swampman does not have genuine beliefs at all. As far as I am concerned, this is a mark against the causal theory of reference, rather than against the genuine content of Swampman's apparent beliefs. For consider: if we asked the Swampman about the city where he remembers having lived, he would tell us informative things (indeed, *true* things – if only they could refer to it!). Someone committed to the causal theory of reference would say that his statements do not express genuine beliefs, they have no genuine propositional content. If that is not strange enough, suppose now that we take the Swampman into the city where he (feels like he) remembers having lived. *Now* there is a causal interaction between Swampman and the city, so the causal theorist will say that what had only appeared to be beliefs before are now real beliefs, complete with referential content. But what would the Swampman say, if we asked him about the city as we arrive? No doubt, among other things, he would say that the city was what he had been talking about all along. To hold that the Swampman's beliefs only gain propositional

content after he's come into contact with the city, as the causal theorist does, strains credulity.

I do not want to turn this into a dissertation on the philosophy of language, but I will take a brief moment to point out that although the causal theory of reference has some strong intuitions in its favour (see, for example, Putnam (1981, ch.1)), it also faces some serious problems of its own. For example, abstract and imaginary objects do not appear to be the kind of objects with which we can have causal interaction – they do not exist in space and time, and we can only have causal interactions with objects that exist in space and time – and yet we can refer to them. The same goes for some objects that are out there in the world: light photons emitted from distant stars, that have not yet reached us, have never come into causal interaction with us, and yet we can refer to the set of all light photons that have not yet reached us. The causal constraint on reference entails that we cannot refer to them, for we have never had any interaction with them. In defense of the causal theory, one might point out that we have had interactions with objects that have had interactions with those photons, and that indirect causal connection is sufficient to establish reference. But that relaxing of the causal theory also allows that Swampman can have real beliefs and a real mind: he is an item in the universe, and there is no doubt some indirect chain of causal interaction linking him with the objects of his beliefs.

The preceding arguments were in support of the claim that Swampman does in fact have a mind, complete with real beliefs with propositional content. What we have to see now is that Swampman's beliefs enjoy the relevant kind of epistemic properties. That

is easy enough to show. There are two points that I want to make. The first is that it is important that Graham admits that Swampman's beliefs enjoy justification. The second is that Graham's modification of the Swampman case is extremely misleading; there is a better way to revise the case, which will show that the Swamp case is a serious problem for the function analysis of entitlement.

The most important kind of epistemic rationality, to my mind, is the kind that goes with reasons for belief. At the very least, that is an important class of rational beliefs. If the proper function analysis of entitlement fails to shed light on justification and reasons for belief, then something crucial is missing from its account. And Graham's account of entitlement is therefore not a genuine candidate for an account of epistemic rationality; his account of entitlement is explicitly divorced from reasons, understood (as he understands them) as warranted beliefs. I do not want to take a stand here on whether the only things that can act as reasons are themselves beliefs. (In fact, I rather doubt it: perceptual input, for example, is cognitively accessible, and so even a strong internalist about justification can appeal to it in an account of justifying reasons, despite the fact that perceptual input need not be in the form of beliefs.) What is important here is that Graham is not attempting to give a proper functionalist account of justification and reasons at all. He admits that Swampman, who has no selected-effects functions of his cognitive systems, can still have justifications; he can cite reasons in support of his beliefs. It is not reasons and rationality, then, that are being grounded in proper functions, but a specifically reason-independent sort of entitlement.

However, I would like to point out that there is still a problem even for the attempt to give a proper-function account of entitlement. Recall that Graham claims that a counterexample to his account of entitlement would have to be a philosophical Swamp-zombie, a being with beliefs and perceptual belief-forming mechanisms but with no rich conscious experience and no justifications for his beliefs. Now, I am not convinced that this is a coherent possibility, but even if it is, it is not the kind of being to which we ought to appeal in constructing a counterexample to the proper-function account of entitlement. Graham notes that the driving force behind externalist accounts of warrant (i.e. positive epistemic status), of which his account of entitlement is a species, is in the first place to be able to account for what is epistemically good in the beliefs of small children and animals, who are not in a position (because they do not have the concepts) to reflect on their beliefs and critically engage with their reasons (2011b, pp.5-6).

The way to construct a counterexample to Graham's proper-function account of entitlement is therefore not to think of a Swamp-zombie, who has no rich phenomenal content. Rather, we need to think of a Swamp-*child* (or a Swamp-dog, cat, monkey, etc.): a molecule-for-molecule replica of a normal small human child, complete with belief-forming processes, and rich phenomenal content. If the ordinary child has entitlement for its beliefs, then so too ought the Swamp-child to have entitlement for its beliefs. Suppose, for example, that you give the Swamp-child a bottle, but it turns out that the bottle is empty. Why does the child cry? Because it believes that the bottle is empty. And that belief surely enjoys every epistemic property that the Swamp-child's doppelganger's

belief about *its* empty bottle does. But, because the Swamp-child has no selected-effect history for its cognitive systems, it does not enjoy entitlement in Graham's sense.

Let's pause here and take stock. We just saw three objections intended to show that making epistemic rationality depend on the natural selection of a type of cognitive system does not work. The Swampman case makes the point that natural selection must not be what grounds either semantic content or epistemic rationality. The other two objections are an attempt to make the same general point in a more mundane way: on the one hand, it is quite possible that our own cognitive systems lack the necessary history to be bearers of proper functions. On the other hand, even if our cognitive systems *were* naturally selected for, still the cognitive systems of early humans, before truth-conducive cognitive systems were selected for, did not have the required histories to ground proper functions – and yet, if such people based their beliefs on good evidence, then their beliefs ought to be candidates for being epistemically rational. The function-account of epistemic rationality, however, is committed to denying their rationality or irrationality.

### **5.3. Proper Functions and Normativity**

The next set of objections to the function-account has to do with the purported normativity of functions. A key feature of the selected-effect account of functions that makes the appeal to such functions so important is that it can handle cases of malfunctioning organs, while still counting such organs as members of the kind of organ that they are.<sup>47</sup> The heart, for example, is the organ whose function it is to pump blood.

---

<sup>47</sup> See Neander (1991b, pp. 466-7). Neander also identifies a second feature of selected-effects accounts of

But some hearts are unable to perform that function. What we want to say in such cases is that the heart is malfunctioning; it has the function of pumping blood, but it is failing to perform that function. We do not, however, want to say that it ceases being a heart (a thing whose function it is to pump blood), when it stops fulfilling the function of a heart. The selected-effects account of proper functions allows us to say that the heart is not just what in fact does pump blood – because some hearts do not do so – but instead, to say that the heart is what is *supposed* to pump blood. Surely, one might say, this “supposed-to” talk indicates some kind of normativity, and so one might want to infer that selected-effect functions can ground the normativity of epistemic rationality.

There are three arguments which show that selected-effect functions cannot ground the kind of normativity required for epistemic rationality.

### **5.3.1. Normativity 1: explanatory vs. normative functions**

The first is that the normativity involved is, at best, only of a very weak sort. As Godfrey-Smith (1996, p.19) emphasizes, “What distinguishes a function from a mere effect is causal/explanatory importance and no more.” We can talk about proper functions as what an item has been selected for, and we can talk about malfunctions as divergences from what an item’s ancestors have been selected for, but there is nothing really normative about what an item’s ancestors have been selected for. For an item to function properly in the proper circumstances is just for it “to do whatever explains why it is there... to malfunction is simply to *fail to do the explanatory thing*. This is a very weak

---

functions as important: they provide a way to talk about functional kinds in biology. Whether or not we need an etiological account of function to do so is not important for us now, however; the important thing is the account of malfunction.

kind of normativity” (ibid.). The point here is that there is no inherent connection between an effect’s *explaining* why an item exists, and that effect’s being *good*, or worthy of promoting, or anything of the sort. When the interests of agents are involved, or when there is some independent reason for which some end or effect is valuable, we can talk about its being valuable or normative. But if all that we have is an explanatory connection between an effect and the existence of an organic structure, then we do not thereby have an end that gives anyone or anything any reason for promoting that end.

To repeat: *all* that is involved in attributing a function to an organ or a trait is to say that its ancestors did that thing, and the fact that its ancestors did that thing is part of the explanation of why the organ or trait exists now and does what it does. We can make judgments of malfunction based on such function-ascriptions, when items fail to do the explanatory thing, but we must not be tempted into thinking that the performance of the function is therefore *good* in any sense. The goodness of fulfilling a function is restricted to the special case where an agent’s interests are involved, or there is some independent reason that makes fulfilling the function a good thing to do. But that is not the general case.

Neander herself makes a similar point, in her entry “Teleological Theories of Mental Content,” in the Stanford Encyclopedia:

Most (if not all) proponents of teleological theories think that functional norms are descriptive and not prescriptive, and the disagreement is over whether it is appropriate to refer to descriptive norms (or “norms”, if you prefer) as normative. Some prefer to reserve the term “normative” for prescriptive contexts, so that a statement would count as normative only if it entails an ought-claim without the addition of further premises. Perhaps most proponents of teleological theories of

mental content would agree that no ought-claim follows from a simple function ascription, not at least without the addition of further premises. (2009, section 2)

Functions do not generate anything normative, in the sense of generating reasons for doing anything, or generating ought-claims. But epistemic reasons are precisely reasons for doing something: they are reasons for holding or rejecting beliefs, or for suspending judgment. And judgments of epistemic rationality are epistemic ought-claims, claims about what people epistemically ought to believe. If function ascriptions do not entail ought-claims, then the ascription of the function of producing true beliefs to our cognitive systems does not entail that we ought to form beliefs in accord with that proper function. So functions cannot ground epistemic reasons, and they cannot ground epistemic rationality.

### **5.3.2. Normativity 2: supernatural functions and normativity**

The second argument I have in mind against the normativity of functions is a related one, but it is targeted at the attempt to ground proper functions in a supernatural designer, as Plantinga (1993) does. Plantinga uses the fact that there is nothing normative in functions understood naturalistically, together with his proper-function account of warrant, as an argument for the existence of a supernatural designer. The objection that I have in mind here is that even if there are functions in nature, and even if they have been put there by the intentional work of an intelligent designer, it does not follow that proper functions give *agents* any reasons for holding or refraining from holding beliefs. Consider the following case. Suppose that there is a designer of the world, and he is a tricky fellow: his purpose in designing hands for us is so that he will get to witness many events of hair-

pulling. Because, in this case, hands have been intentionally designed for a purpose, they satisfy the conditions for having a proper function. But it does not follow that I now have any reason to pull anyone's hair – indeed, that does not even follow if we add the additional premise that I *know* that my designer had this purpose in mind. I recognize that pulling people's hair is not a good purpose, so I have no reason to do it. Intelligent design of nature, and in particular of hair and hands (and cognitive systems), is not sufficient to generate reasons for agents to try to ensure that hair and hands (and cognitive systems) fulfill their proper function.

Plantinga would respond, no doubt, that this argument is irrelevant, because his position is that there must be a *good* designer, if there are to be normatively binding proper functions in nature. However, just as we do not want to hang the epistemic rationality of our beliefs on the uncertain possibility that our cognitive systems have the right kind of history of natural selection, so too we do not want to hang the epistemic rationality of our beliefs on the uncertain possibility that our cognitive systems have been designed by a good intelligent designer. If we did hold that epistemic rationality was dependent on such a designer, then in the event that there is no God, or that there is a God but that he did not design our cognitive system, epistemic rationality would not apply to us.

Someone might object: “But Plantinga's concept of God is that of the necessary being, the one that exists at *every* possible world! (see Plantinga 1974). So it is not uncertain whether God exists.” My reply to this objection is, first, that the existence of

Plantinga's God is metaphysically necessary, if it is true at all, but (as Plantinga admits: 1974, p.221), it is still epistemically uncertain, and that is what matters for us now: we do not know whether such a God exists. We do know, however, that there are epistemically rational and irrational beliefs, so we should not hang our analysis of epistemic rationality on the existence of God.

Second, and more importantly, even if God exists at every possible world, it is not the case that God *designs* us at every possible world; there are possible worlds in which we evolve by natural selection, without any planning on God's part. At those worlds where God does not design us, but simply allows us to evolve, our cognitive systems would lack normatively binding proper functions. But still, at those worlds, our beliefs could be epistemically rational or irrational. (Indeed, *this* could be such a world.) So we cannot hang the analysis of epistemic rationality on the intentions of an intelligent designer.

### **5.3.3. Normativity 3: the case of the rape gene**

Finally, the third argument that I have in mind against the normativity of proper functions is an argument from example, designed to drive home the point that neither natural selection nor intelligent design is sufficient to make it good for proper functions to be fulfilled, and therefore that proper functions do not by themselves give rise to reasons to do anything. Perhaps I am being a bit repetitive, but I want to emphasize the point that proper functions do not themselves give people reasons to do anything.

The example is as follows. Suppose that there is a population of people, where most of the people have a certain gene, which causes in the men a strong desire to commit rape, and it causes in the women an extreme fear of resisting when they are assaulted. Because of this desire, very many men have in fact committed rape in the past, and many children have been born as a result, and the gene is now possessed by most members of the population. So, if we want to say (as most of us do) that the gene (or cluster of genes) that handles little toes in humans has the proper function of giving rise to little toes in humans, because it was naturally selected for that function, then we also must say that the rape-gene has the proper function of giving rise to rape, because it was naturally selected for that function. But, quite obviously, the fact that the gene has that proper function gives the men no *reason* whatsoever to commit rape; what men ought to do is to *resist* that desire. By analogy, the fact, if it is a fact, that our cognitive systems have been naturally selected for the function of forming true beliefs does not by itself give us any reason to go ahead and form true beliefs.

The example works even if we suppose that God designed the rape-gene in order to give rise to the desire in men. For, supposing that God designed the world, we know that there are many things that he designed that we ought not to promote. According to the beliefs of many religious people, for example, our animal desires are given to us precisely to be *overcome* rather than wantonly satisfied.

To repeat Godfrey-Smith's point, then, proper functions, being merely causal-explanatory things, are not normative things. They cannot give rise to normative reasons

to do anything, including to form or to refrain from forming beliefs, or any kind of epistemic ought-claim. Proper functions therefore cannot ground epistemic rationality.

## **Chapter 6: Instrumental Rationality is Not Sufficient for Epistemic**

### **Rationality**

In the fifth chapter, I began the task of arguing against the instrumental conception. That chapter was concerned with the proper functionalist account of epistemic rationality, which is plausible to take as a version of the objective instrumental conception, because it holds that the epistemic goal is determined by the proper function of our cognitive systems, whether or not we care about it. For the purpose of the arguments I give in the following chapters, I assume that the proper function account of epistemic rationality is not viable, so I do not discuss it any further. In the present chapter, I continue the case against the instrumental conception, by presenting an objection to the effect that believing in such a way as to in fact achieve, or else in such a way that it is reasonable to think that one will achieve, the epistemic goal, is not sufficient for epistemic rationality. It follows, then, that instrumental rationality in the service of the epistemic goal is not sufficient for epistemic rationality.

### **6.1. The Problem**

The purpose of this section is to set up a case where a subject forms a belief which is appropriately directed to the achieving of the epistemic goal but is still an irrational belief. The point of the example is to show that achieving the epistemic goal in the right way is not sufficient for epistemic rationality. The objection is not conclusive, as there is a way to revise the conception of the achievement of the epistemic goal that avoids the problem, but the revision brings its own problems, which we'll get to shortly.

First, the case. It is an extended case, but not hard to follow.

### **Omar**

Omar is a college student who has been told his whole life that he is an infallible hornbeam identifier. As a matter of fact, he is extremely reliable, but not infallible; he has occasionally identified trees as hornbeams when they were not. He has even, on two occasions, recognized his mistake. However, he has put those occasions from his mind; he never thinks about them. He dwells instead on the constant reinforcement of the belief in his infallibility provided to him both by people around him who tell him that he is infallible, as well as by the many instances of successful hornbeam identification that he remembers. If Omar were to reflect on his evidence, he would quickly recognize that he is not an infallible hornbeam identifier.

Omar has recently taken an epistemology course, and he is now impressed by Richard Foley's instrumental conception of epistemic rationality. He therefore takes it upon himself to do his best to ensure that his beliefs achieve the epistemic goal of now believing true propositions and now not believing false propositions. He begins with the belief in his infallibility as a hornbeam identifier.

Here is the twist. Omar has also recently taken an anthropology class on tree-identifiers. He has learned in that class that there are scientific studies which show that people who believe themselves to be infallible identifiers of tree-types are in fact more reliable in their identifications than people who do not take

themselves to be infallible tree-identifiers, and they form many more beliefs about trees. Moreover, when people who had previously taken themselves to be reliable tree-identifiers give up the belief in their infallibility, they invariably at the same time lose a whole host of their beliefs about trees and about themselves as epistemic agents.

Omar then reflects: “Well, if I believe that I am infallible, then even if I am mistaken in that belief, it still promotes the achievement of the epistemic goal. Because I believe myself to be infallible, this study shows that whether or not my belief is true, it sustains many other true beliefs that I have. If I give up the belief in my infallibility, then I lose those beliefs too. But the epistemic goal is now to believe those propositions that are true, and now not to believe those propositions that are false. Whether my belief in my infallibility is true or false, it causally sustains many true beliefs now. So I am epistemically rational in retaining that belief. Good!”

Because Omar has recognized his mistaken identifications before, and he could easily recall those occasions if he thought about it, his belief in his own infallibility ought not to count as epistemically rational.<sup>48</sup> But if epistemic rationality is a matter of achieving the

---

<sup>48</sup> Recall that it does not beg the question against the instrumentalist here to hold that beliefs held contrary to readily available evidence are epistemically irrational. As I argued in Chapter 1, we can appeal to common sense intuitions about concepts and cases as a constraint on our theorizing about our concepts. What I am up to here is appealing to what should be an intuitively obvious case of epistemic irrationality, which the theories need to be able to account for. An account of epistemic rationality must not deliver the wrong verdict in the cases of obvious epistemic rationality or irrationality. Absent positive reason to think that we are unreliable in a given case, in analyzing our concepts, we may legitimately employ natural, widely-shared, unforced intuitions about whether cases fall under a concept, as a constraint on our theorizing about that concept. Furthermore, recall (from Chapter 4) that instrumentalists also want to

epistemic goal, then his belief does count as epistemically rational. That is true whether we take the actual achievement of the epistemic goal as our standard of epistemic rationality, or only whether the subject must think (or have good reason to think, or would think upon sufficient reflection) that the belief achieves the goal. For Omar's belief does in fact promote the achievement of the epistemic goal of now believing truths and now not believing falsehoods: it causally sustains many true beliefs that would be lost immediately upon giving up the belief, and it is only one false belief. Furthermore, Omar in fact takes his belief to promote the achievement of the epistemic goal, and he has good reason to think that it does so, and he would think so upon reflection (indeed, he has reflected seriously upon it). After all, he has read a reputable study that indicates that his belief in his infallibility is causally sustaining many other true beliefs that he holds.

This is bad news for the instrumental conception. This objection applies to both the subjective and the objective instrumental conception, because whether the goal is one that Omar cares about achieving (which he does) or whether the goal is independently valuable and to-be-promoted, Omar is instrumentally rational in his attempt to achieve that goal. Yet his belief ought not to count as epistemically rational.

## **6.2. Objection 1: No Doxastic Voluntarism**

There are three ways to try to avoid this result. First of all, one might object that Omar really does not believe that he is infallible. If he has recognized his mistaken

---

respect the importance of evidence for an account of epistemic rationality, so it does not beg the question against them to appeal to cases of belief based on good evidence as cases of epistemically rational belief.

identifications in the past, then maybe he would *like* to continue to believe in his infallibility, but he cannot.

The force of this objection comes from its denial of doxastic voluntarism, a simplified version of which is that we can believe what we like, even when the evidence is clearly against our beliefs. I will not undertake a full discussion of doxastic voluntarism, but I would like to say that the objection underestimates how clever we humans can sometimes be at getting ourselves to believe things against the evidence and against what we remember. The Kierkegaardian view of faith, for example, as belief that flies in the face of the evidence, is something that at least one person presumably took seriously (I mean, of course, Kierkegaard himself). Or take the more mundane case, say, of the beliefs of baseball players. Even the best batters only bat around .350, which is to say that 65% of their at-bats do not result in a hit. Still, batters go up to the plate, thinking “*this* time, I’m going to hit the ball.” They think that they will get a hit, even though they know that the odds are against them. In other words, they recognize that the probability of their getting a hit is low, but they believe that they will get a hit anyway. That is not to say that there is no good reason for them to think that they will get a hit: players who believe that they will get a hit manage to get hits more often than players who do not believe in themselves. That is a very good reason for them to believe that they will get a hit, each time they go up to bat, but that reason is practical rather than epistemic.

In any case, the objection from the rejection of doxastic voluntarism really seems most troubling only if we think of Omar as *simultaneously* entertaining the thoughts that

he is infallible and that he has been mistaken in the past. But just as a batter need not be holding his batting average before his mind when he goes up to the plate thinking he will get a hit, so too Omar need not be holding the mistaken identifications before his mind when he thinks about his infallibility.

### **6.3. Objection 2: Rule-Consequentialism**

The second way to try to avoid the result is to go rule-consequentialist, and make epistemic rationality depend on the *rule* or set of rules for belief formation that best achieves the epistemic goal. Goldman (1986, p.97), for one, is explicitly interested in rule-consequentialism, in the service of the goal of believing truths, as underlying the framework of epistemic justification.<sup>49</sup> The idea is that the consequences of Omar's belief in his infallibility do not count toward its epistemic status; it is the *process* by which the belief was formed that justifies it. The process is licensed by a framework of justification-rules, which are themselves truth-directed. Processes confer justification, on this picture, if they do well vis-a-vis the truth-goal.

However, the rule-consequentialist proposal does not offer a satisfactory solution to the case of Omar. Omar is in fact forming his belief in accord with an excellent rule vis-a-vis the epistemic goal: believe only those propositions that you have good reason to think will achieve the epistemic goal of now believing those propositions that are true and now not believing those propositions that are false. Any system of truth-directed justification-rules will have to license that process of belief-formation, because (1) it does

---

<sup>49</sup> Leite (2007) suggests a similar move for instrumentalists.

in fact do well vis-a-vis the truth-goal, and (2) Omar believes on reflection that it does well vis-a-vis the truth-goal. So, whatever else we can say about the case of Omar, it presents a problem for a rule-consequentialist approach to instrumental epistemic rationality.

#### **6.4. Objection 3: The Myopic Restriction**

And so we come to the third way of trying to avoid the consequence that the instrumentalist is committed to saying that Omar's belief in his infallibility is epistemically rational. This objection is to hold that the epistemic goal is myopic, in the sense that beliefs must each *in and of themselves* achieve the epistemic goal; epistemic evaluations in light of the epistemic goal do not "see through" to a belief's causal consequences for adopting, retaining, or losing other beliefs. After all, Foley's main concern in formulating the epistemic goal as a present-tense goal is to eliminate the possibility of having obvious cases of epistemically irrational beliefs count as achieving the epistemic goal, by way of their causal consequences for the future adoption of beliefs. It is in that same spirit of trying to get the extension of epistemically rational beliefs right, that we can require that each and every belief achieve the epistemic goal, ignoring even the present-tense consequences that adopting or rejecting a given belief will have for one's broader body of beliefs. Marian David makes that point about the reason for the present-tense restriction on the epistemic goal explicit:

Being justified in believing  $p$  has nothing to do with the causal consequences of believing  $p$ . More generally, it seems that being justified in believing  $p$  has nothing to do with what beliefs you are going to hold in the future. The truth-goal cannot be a diachronic goal if it is to play the role assigned to it in the goal-

oriented approach to justification... if it were, the causal consequences of our beliefs would be relevant to their epistemic status. Instead, it must be a synchronic goal: it must be the goal of *now* having beliefs that are true and *now* not having beliefs that are false. (2001, pp.160-161)

The reason that we have restricted the epistemic goal to the present tense was in order to avoid counting irrelevant factors toward the epistemic rationality of beliefs. The present suggestion, of restricting the evaluation of beliefs to how well they fare, in and of themselves, in relation to the epistemic goal, is just an extension of that same move.

This way of handling the objection does solve the problem at hand. If we require of each belief that it promote the achievement of the epistemic goal all by itself, then Omar's belief no longer counts as epistemically rational, since it only achieves the epistemic goal by virtue of its causal relations with other beliefs. However, this restriction on the achievement of the epistemic goal brings with it a number of problems of its own. The problems are as follows.

(1) For one thing, putting in the restriction that each belief in and of itself must do well vis-a-vis the goal is an ad hoc restriction, which lacks good motivation from an instrumental point of view. In general, given some good G, it is those means that will promote more of G than those that promote less of G that are best instrumentally justified relative to that good (absent independent reasons for preferring less of G). That is, actions are not generally evaluated for instrumental rationality insofar as they promote the achievement of a goal in and of themselves – for one thing, actions are typically instrumentally evaluated with an eye to their many effects; and for another, their success

at achieving the goal in question is weighed against other possible actions and how well *they* would succeed.

To illustrate that point, consider the following case. Tory is going camping at the end of the week, and she will be gone for a month. Before she goes, she wants to see some movies. Her parents, with a view to teaching her responsibility, have given her a choice. She can see two movies, if she does her chores tonight, or she can see one movie tonight, but then she gets to see no more movies before she goes camping. (To simplify the case, suppose that Tory does not much mind doing chores, and that she will not get tired of movies after seeing one this week.) Tory's options, then, are to see a movie tonight, or to do her chores tonight so that she can see two movies in the following days. Vis-a-vis her goal of seeing movies, it is practically rational for her to do her chores tonight, because she will then be able to see two movies. However, considered in and of itself, irrespective of its causal consequences, the action of seeing a movie tonight promotes her goal of seeing movies, and so that comes out as practically rational. Now, perhaps we should say that it is practically *permissible* for Tory to choose to see a movie tonight (since it does promote the achievement of her goal), but surely it is practically *better, more* practically rational, for her to do her chores tonight and see two movies later.

All of that is just to say that instrumental rationality is not in general limited by the "in-and-of-itself" restriction that we are considering with respect to epistemic rationality. Adding in the restriction when we turn to the epistemic case appears to be unmotivated, from an instrumental point of view, except in order to get around an

objection. Instrumental rationality can in general be transferred to a belief or action by the fact that it promotes the achievement of a goal – and if the belief or action in question promotes having more beliefs or performing more actions that themselves promote the achievement of the goal, so much the better, so far as the instrumental rationality of the initial belief or action goes.

Of course the ad hoc nature of a response to an objection is not always a damning problem, but other things being equal, it is best to be able to respond to objections with claims that are well-motivated from the point of view of the position being defended.

(2) Another small problem that comes with the restriction that each belief in and of itself achieve the epistemic goal is that I am aware of very few instrumentalists who advocate such a restriction. Certainly no rule-consequentialist will want one, and I am not certain that there is anyone else who explicitly advocates one either. The only person I am aware of who looks like he might be explicitly claiming that beliefs face the epistemic goal individually is Tannsjö: “I compare each of the beliefs I adopt... with its negation, and with the possibility, in relation to each of them, of suspending my judgement. These comparisons, and these comparisons only, are relevant from the point of view of my epistemic goal. The epistemic goal, then, has built into it a kind of myopia” (2010, p.111). If Tannsjö means to say that beliefs face the epistemic goal without regard to any of their effects for the belief-system as a whole, then he is advocating the kind of restriction that I have in mind.<sup>50</sup> But then Tannsjö swiftly proceeds to go back on the claim that the

---

<sup>50</sup> It is not clear that that is his point, though: Tannsjö makes his claim about how beliefs face the epistemic

epistemic goal is myopic in this sense: “This myopia must be given up, however, when we consider the epistemic goal and tries [sic] to decide what weight to give to the two goals, of believing what is true, and not to believe what is false, in our application of them” (ibid., p.112). Because we have to weigh how much relative weight to give to the two aspects of the epistemic goal, we are interested in determining how well a *body* of beliefs will do with respect to the epistemic goal, and so the effects of a belief for the whole of a body of beliefs become relevant.

My worry, then, is that I do not want to be putting words in the mouths of instrumentalists that they have not spoken. But, because the restriction that beliefs must face the epistemic goal in and of themselves seems to be required to deal with cases like Omar’s, it is perhaps the charitable thing to do to build it into the position.

(3) A more serious problem with the restriction that each belief in and of itself come up for evaluation in light of the epistemic goal is a modified version of Maitzen’s objection to the truth-goal, which we saw in chapters 2 and 3: all true beliefs turn out to be epistemically rational, and no false beliefs turn out to be epistemically rational. This follows, given Giere’s objective sense of instrumental rationality that we saw in chapter

---

goal, in the context of a discussion of a comparison between different fields of study. In physics, for example, you could acquire many more true beliefs than in philosophy; but when you are doing philosophy, your candidate beliefs are not weighed against the candidate beliefs that you would have gained had you gone into physics rather than philosophy; you weigh your candidate philosophy-beliefs only against suspending judgment or believing the negation of those same beliefs.

It is compatible with this restriction on how beliefs face the epistemic goal – as weighed only against positive disbelief or suspension of judgment – to say that the consequences that a belief has for the belief-system overall factor into the decision whether to believe, disbelieve, or suspend judgment. So it is not clear to me just what Tannsjo means here.

2. And a similar problem will also follow, given Giere's subjective sense of instrumental rationality.<sup>51</sup>

Recall that objective rationality, for Giere, is a matter of taking the means that will in fact achieve the goal in question. Subjective rationality is a matter of taking the means that a subject believes will achieve the goal. Now, all true beliefs are means that will in fact achieve the epistemic goal, and no false belief will do so, if beliefs come up for epistemic evaluation in light of the epistemic goal individually, and the goal is now to have all and only true beliefs. It follows that, on Giere's objective sense of instrumental rationality, all and only true beliefs are epistemically rational.

On the other hand, all beliefs are such that the subject who has them takes them to be true (because to believe that *p* is just to take *p* to be true). So if instrumental rationality is subjective in Giere's sense, then all and only those beliefs that a subject actually has will come out as epistemically rational for that subject, because they are all such that the subject who has them (implicitly) takes them to be appropriate means for achieving the epistemic goal (or at least, she would take them to be, if she took an epistemology course, and became aware of what the epistemic goal is). Neither of these results is acceptable: it is not the case that all and only true beliefs are epistemically rational, nor is it the case that all and only those beliefs of hers that a subject thinks are true are epistemically rational for her.

---

<sup>51</sup> I should note that Giere is not the only philosopher to talk about objective and subjective rationality in these ways; it is quite common to talk this way. But Giere is clear and straightforward about it, and he is himself a prominent representative of objective instrumentalism in this sense of the term, and subjective instrumentalism in my sense of the term, so I refer to his sense of the terms "subjective" and "objective" to distinguish them from mine.

Foley's conception of rationality, as what an agent would *upon sufficient reflection* take to achieve a goal, survives that argument, because it is not necessarily the case that an agent would retain all of the beliefs that she in fact has, if she were to reflect sufficiently. There is room on Foley's account for subjectively irrational beliefs, even given the in-and-of-itself restriction on the way that beliefs face the epistemic goal. Foley is rather alone among contemporary epistemologists in adopting the sufficient-reflection conception of rationality, though, and for everyone who adopts either Giere's subjective or his objective senses of rationality, the unacceptable consequences will hold.

### **6.5. Sufficient Reflection and Properly Based Beliefs**

So at this point, I want to offer an argument against Foley's sufficient-reflection conception of epistemic rationality. Until now I was content to let it pass as just one more way to be an instrumentalist, but now it is important to be able to show that there are problems specific to the sufficient-reflection view of rationality.<sup>52</sup> The sufficient-reflection view that Foley advocates is that it is epistemically rational for a subject to believe those propositions that she would, upon sufficient reflection, take to achieve the epistemic goal. Sufficient reflection is reflection with only the epistemic goal in mind, ignoring all other goals and distractions, and it continues up until the point of reflective stability: the point where, if she were to continue reflecting, a subject would not change her mind (1987, p.35; 1993, p.99). When reflection reaches this point, it reflects a subject's deepest epistemic standards.

---

<sup>52</sup> It will also be important to be able to show that Foley's sufficient-reflection conception of rationality is problematic when we get to the regress argument in Chapter 8.

The main problem with the sufficient-reflection conception of epistemic rationality is that it completely divorces epistemic rationality from the reason for which a belief is held. A person might hold a belief, and she might have good evidence for the belief, but fail to see that she has good evidence for it. Instead, she might hold the belief for some poor reason. Foley's view is that that belief is nevertheless epistemically rational.

Let's consider some cases.

### **True Mathematical Belief**

Suppose someone I trust has told me the answer to a complicated mathematical problem. Suppose that I do not now see the calculation, but I take his word for it. I am capable enough at mathematics, and if I thought about the problem for a very long time, I could see the calculation for myself. So my belief, based on what my friend has told me, conforms to my deepest epistemic standards.

### **False Mathematical Belief**

Suppose again that a friend whom I trust has told me the answer to a complicated mathematical problem, and again I do not see the calculation, but this time my friend has made a subtle mistake. Again, if I thought about it for a very long time, I could figure out the calculation myself, and I would be convinced that he was wrong. So the belief does not conform to my deepest epistemic standards.

In the True Mathematical Belief case, my belief is epistemically rational, on Foley's account, whereas in False Mathematical Belief, it is not. But in both cases, my belief is

held for very good reason (the testimony of a source that I know to be reliable), so what we should say is that in both cases, my belief is epistemically rational.

Now consider another case.

### **Skeptical Peter**

Peter is an amateur philosopher. He has taken an interest in skepticism about the external world. He has not carefully considered the arguments for or against skepticism, to determine whether they are correct, but he likes how easy it is to use skeptical arguments to fluster other amateur philosophers. Because he has so much fun using skeptical arguments, he has come to believe them. People have sometimes given good arguments against his views, and he has recognized that they are very good arguments, which he could not answer on the spot or afterwards in thinking about it, but he has not bothered pursuing the issue very deeply.

As it happens, if Peter were to be sufficiently reflective, he would, after long deliberation, determine that the skeptical arguments that he knows of really are decisive. But he has not reflected anywhere close to sufficiently on the matter; he has only come to believe the skeptical arguments because they are fun to give, and because he has repeated them so many times.

In this case, Peter's skeptical beliefs appear to be epistemically irrational, because they are held for poor reasons (how much fun one finds an argument to be is a poor indicator of how good the argument is). Furthermore, he holds his view in the face of good

arguments that he does not currently know how to answer. At the very least, there is something glaringly bad in Peter's beliefs, from an epistemic point of view. If we nevertheless want to say that his beliefs count as epistemically rational, then epistemic rationality is a pretty poor term of epistemic appraisal. A proper basing requirement would handle this case, though: if the considerations that Peter would see upon sufficient reflection to support his skeptical views are not the reasons *for which* he holds his skeptical beliefs, then his beliefs are not epistemically rational.<sup>53</sup>

Foley considers the possibility of causal requirements on rational beliefs in his (1987).<sup>54</sup> Two important reasons he gives for resisting such a requirement are as follows. First, he thinks that there are serious problems for an account that introduces a causal requirement on rational beliefs: (1) any such account must deal with causal deviance, and (2) it must be able to give an account of *how much* of the available evidence for a proposition must play a causal role in the formation or sustaining of the belief. These are hard problems.

The second reason Foley gives for resisting a causal basing requirement for epistemic rationality is that he claims to be interested primarily in those *propositions* that are rational *for* an individual to believe, from the perspective of what a person would

---

<sup>53</sup> See Feldman (1989) and Alston (1989) for some similar types of cases.

<sup>54</sup> I move indifferently between talking about causal requirements and basing requirements here, because when Foley talks about these issues, he talks in terms of the "causal-historical" sense of rationality. The basic idea is that of a basing requirement – of being the reason (or one of the reasons) *for which* a belief is held – and I am open to different ways in which a belief might be based on the reasons for it. In particular, a belief might be based on reasons when it is either directly caused by them, or else directly caused by something entirely unrelated to those reasons, *but*, were the unrelated factors absent, the reasons in question *would* then directly cause the belief. There is plenty to say here about the basing requirement, but I do not propose to get into it any further; the important thing is just to point out that I am not intending to take sides on what kind of basing relation is necessary when I talk about a causal requirement.

believe upon sufficient reflection. His view, he writes, is not primarily about epistemically rational *beliefs* (1987, p.179). The distinction here is between beliefs for which one merely *has* evidence or reasons, on the one hand, and beliefs for which one has evidence or reasons, where the beliefs are also appropriately caused or sustained by the evidence that one has for them. The first kind of rationality is the rationality of propositions (sometimes called *propositional* rationality or justification); the second, the rationality of beliefs (sometimes called *doxastic* rationality or justification). The idea is that for a belief to be doxastically rational, it must be based in the right way on the evidence for it. The distinction may also be put in terms of whether the *content* of a subject's belief is epistemically rational for him, on the one hand, and whether his *having* a belief with that content is epistemically rational, on the other. Or the distinction may be put as that between beliefs that *could* be justified for a subject, and those that *are* justified for a subject.<sup>55</sup>

I pointed out in Chapter 1 that in order for a belief to be epistemically rational, I take it that it must meet a basing requirement of some sort (though I was not particular

---

<sup>55</sup> One might argue that these different ways to draw the distinction are not all equivalent. I think that's correct. In fact, I am not sure that there is an interesting distinction to draw between rational propositions to believe and beliefs that are rational. The most we should say, I believe, is that there are some propositions that one *could* rationally believe if one based them on evidence or reasons one currently has, even if one does not rationally believe them now.

However, the basing relation, as well as the ways to distinguish propositional and doxastic justification, needs a good deal of work. I would point, as a good place to start rethinking these matters, to Turri's (2010) and (2011). Turri reverses the ordinary order of explanation: in his analysis, doxastic rationality is employed to explain propositional rationality, rather than the reverse. The usual way to think of the relation between the two is to give an account of propositional rationality or justification, and then to say that a belief is doxastically rational if there is propositional justification for it, and the belief is held on the basis of whatever it is that provides the propositional justification.

about what kind of basing requirement it must meet).<sup>56</sup> Foley's view here, on the other hand, is that propositional rationality is the more interesting kind of rationality.

I think there are convincing replies to be made on behalf of a causal basing requirement on epistemic rationality. In reply to the objection that causal deviance is a problem, we can simply say that it is only non-deviant causation that we mean. That sounds like a cheeky response, I admit, but after all, the basing requirement is only intended to be a necessary condition for epistemic rationality, not a sufficient one. So the mere fact that a belief is caused by good reasons need not always entail that the belief is epistemically rational; if the belief is based on good reasons but caused in an improper way, then the belief is importantly lacking – this is the causal deviance problem. (John Turri (2010) gives a nice analogy: when building a deck, it is not enough to have good lumber and tools on hand; it is not even enough that you *use* those tools and that lumber in building the deck. In order to have a good deck, it must be made *in the right way* from those materials.)

Causal deviance is a problem, and it indicates that we should not think that just any causal relation between reasons and the beliefs that they cause is sufficient for properly basing the beliefs on those reasons. But causal deviance certainly does not give us any reason to think that a causal basing relation is not *necessary* for a belief to be epistemically rational. (In order to build a good deck, one might say, although it is not sufficient for building a good wooden deck that you use good lumber and good tools, it is

---

<sup>56</sup> Requiring that beliefs be caused in the right way by what justifies them is not an idiosyncratic quirk of mine; it appears to be the majority view among epistemologists.

still *necessary* to use good lumber and tools.) So although a complete account of epistemic rationality would require an adequate treatment of causal deviance, we do not need to worry about it for now, because the claim here is only that a causal basing relation is necessary. (I do claim that *proper* basing is sufficient for epistemic rationality: for example, in the True and the False Mathematical Belief above, it looks like the beliefs are epistemically rational because they are held on the basis of a good reason. But in these cases, there is nothing causally deviant to worry about.)

Similarly, we do not need to worry about just how much of one's evidence a belief must be based on, in order for the belief to be doxastically rational. For again, the important part of the proposal here about a causal basing requirement is that it is a *necessary* condition for epistemic rationality. So we can just say that, at a minimum, S's belief that *p* must be based on at least some of S's evidence or reasons for *p*, if *p* is to be epistemically rational for S.

I turn now to Foley's other reason for resisting a causal/basing requirement: that his account is primarily about what propositions are epistemically rational to believe, rather than about epistemically rational beliefs (i.e. about propositional rather than doxastic rationality). There are two replies to make. First: it is important that the account has some consequences that are very difficult to reconcile ourselves with, e.g. in the False Mathematical Belief case above. The beliefs in that case and in its True counterpart are based on a very good reason (indeed, the very same reason), and both seem therefore to be epistemically rational. But Foley's view entails that the belief in False Mathematical

Belief is epistemically irrational. More precisely: Foley's view entails that the belief in that case is propositionally irrational, and the standard view is that a belief must at least be propositionally rational if it is to be doxastically rational (when the belief is held on the basis of what gives it its propositional justification). So Foley's view entails that the belief is epistemically irrational in the doxastic sense as well.

One might object that I am just intuition-mongering, without saying anything useful about why I think it is wrong to focus on propositional rather than doxastic rationality. Well, I *am* seeking to generate intuitions by thinking up possible cases, but I don't think that's a bad thing. As I have said, we ought to be able to appeal to natural intuitions in straightforward cases as constraints on our theorizing. And I think that the intuitions here do indicate that there is a basing requirement on epistemically rational beliefs.

But we do not have to rest with cases and the intuitions that fall out of them, because Foley does not appear to be exclusively interested in giving an account of propositional rationality anyway. It looks like he must have an account of doxastic rationality in mind in at least some places; this is especially clear when we consider his account of *justified* belief. Foley develops an account of justified belief, which is closely related to, but distinct from, his account of epistemically rational belief. Justified belief, in Foley's sense, is sensitive to memory and time resources, as well as to the importance of the issue at hand. It is sensitive in these ways, because this concept is meant to track our ordinary judgments of epistemic responsibility. Foley writes: "one justifiably believes

a proposition P if one has an epistemically rational belief that one's procedures with respect to P have been acceptable; that is, acceptable given the limitations on one's time and capacities and given all of one's goals" (2005, p.322).

The first thing to note about that characterization of justified belief is that it is only given as a sufficient condition: *if* one has an epistemically rational belief that one has done an acceptable job in one's procedures with respect to P, *then* one's belief that P is justified, in this responsibility-tracking sense. So the account is open to justified beliefs where the subject does not have a higher-order epistemically rational belief about the initial belief. Perhaps it is natural to think that what Foley gives as a sufficient condition might also be a necessary condition on justified belief, though. If so, then we might take Foley's sense of justified belief to be captured in the biconditional: "one justifiably believes that P if and only if it is epistemically rational for one to believe that one's procedures with respect to P have been acceptable."

But we do not have to worry about necessary conditions for now. The point is that, in order for the sufficient condition for justified belief that Foley gives us to be at all plausible, the epistemically rational belief *about* the first-order belief to which he appeals must be rational in the *doxastic* sense (it must be *held for* good reasons). Why must it be doxastically rational? Because the concept of justified belief is supposed to track judgments of epistemic responsibility, and it cannot do that job if epistemic rationality is understood in the sense of what a subject *would* believe upon sufficient reflection. For suppose that a subject S has a belief P, and she also has a higher-order belief Q, to the

effect that she has done an acceptable job with respect to the way that she has formed the belief P. Suppose also that Q is propositionally rational for S, in Foley's sense of being such that, *if* she were sufficiently reflective, then she would take Q to achieve the epistemic goal. But suppose that S is *not* sufficiently reflective; she believes Q for a bad reason (say, she read it in the stars).<sup>57</sup> Then S is not responsible in believing that P. But, given the sufficient condition for justified belief above, S's belief that P *is* justified. So justification in this sense does not track responsible belief, if we do not require epistemically rational beliefs to be properly based.

However, if we *do* impose a basing requirement on epistemically rational belief, then it is much more plausible to think that justified belief in this sense will turn out to track judgments of responsible belief. If there are good reasons available for believing Q, and S believes Q for those reasons, then it looks like Q is capable of making P justified for S, if Q is the belief that P will do an acceptable job of achieving S's goals. Since a basing requirement provides just the fix needed to make the account of justified belief do what Foley intends it to do – track judgments of epistemic responsibility – the charitable interpretation is that that is what Foley had in mind all along. Or, if he did not have that in mind all along, then the fact that his account of justified belief only makes sense given a basing requirement on epistemically rational belief gives us a reason to think that Foley needs to either revise his view of epistemic rationality, to incorporate a basing

---

<sup>57</sup> Assume that, had S not had this bad reason, she would not then believe P on the basis of the good reasons available, so that there is no causal overdetermination involved in her belief. Causal overdetermination can plausibly be built into the basing relation, as a counterfactual type of causation, and the point here is to give a clear case where the belief is not properly based on good reasons.

requirement, or else revise his view of justified belief, and give up on the idea of trying to track epistemic responsibility.

A tempting move here might be to insert an explicit basing relation into Foley's account, so that a proposition  $p$  is *propositionally* rational for a person if it is what she would believe upon sufficient reflection, and  $p$  is *doxastically* rational (it is a rational belief, it is held for good reason) if she *has* reflected from the epistemic point of view up to the point of reflective equilibrium. But that move faces the problem that very many of our beliefs that appear to be epistemically rational, given that they are based on good evidence, will turn out to be epistemically irrational, since (I take it) we do not often in fact reflect from the epistemic point of view to the point of reflective equilibrium. Many of the beliefs that we have, even if they are based on good evidence, will be such that long, sustained reflection could change our minds about them. We might change our minds about them many times; we might even come back to our initial belief. But as long as thinking about it some more would change our minds, then the belief is not propositionally rational on Foley's account, and so it is not doxastically rational.

All of that was to argue against Foley's sufficient-reflection notion of epistemic rationality. Now I want to give a possible reason why beliefs that are improperly based appear to be epistemically irrational, while those that are based on good reasons appear to be epistemically rational. The proposal is that it is because beliefs that are properly based on good reasons are (basis-relative) *safe*, while beliefs that are improperly based, however much propositional rationality they might have, are (basis-relative) *unsafe*. The basic

notion of basis-relative safety is this: S's belief that  $p$  is basis-relative safe if, given the basis on which  $p$  is held,  $p$  could not easily be mistaken. S's belief that  $p$  is basis-relative unsafe if, given the basis on which  $p$  is held,  $p$  could easily be mistaken. In other words, your belief that  $p$ , which you hold for reason  $q$ , is *safe*, provided that in most close possible worlds in which you believe  $p$  for reason  $q$ ,  $p$  is true.<sup>58</sup> Basis-relative safety comes into play here, I suspect, because when a belief is held for a reason, it ought not to be the case that the belief could be held for that reason and be easily mistaken. When a belief is basis-relative safe, it is based on a good reason, and it is epistemically rational.

In fact, not only is an improperly-based belief unsafe with respect to its *truth*; it is also unsafe with respect to whether it is propositionally rational. If a subject holds a belief on a poor basis, say for reason R, then whether or not there is in fact some sort of propositional justification for the belief that is available, still it is the case that there are close possible worlds in which the belief is held for reason R but in which there is no available propositional justification. This is guaranteed by the fact that propositional justification has nothing whatsoever to do with the actual basis on which beliefs are held.

These considerations about safety are largely speculative, and nothing big hangs on them. Still, I thought it appropriate to say a few words about why it might be that the basing relation seems to many of us be so important, and this proposal seems at least plausible to me.

---

<sup>58</sup> Sosa discusses the notion of safety, for example, in his (1999) and (2007).

## **Chapter 7: The Value of the Epistemic Goal**

The basic argument of this chapter is that in order for it to be instrumentally rational to take the means to achieve a goal, that goal must at least be in some minimal sense good to achieve, either because it is in fact desired by some agent, or because it is good to achieve without reference to the agent's desires. But there are cases where the epistemic goal is not good to achieve in any plausible sense, and yet subjects can have epistemically rational or irrational beliefs in such cases. The conclusion to draw is that epistemic rationality is not instrumental.

### **7.1. Movie-Goers, Unfaithful Spouses, and Grains of Sand**

Recall that there are two ways to be an instrumentalist about epistemic rationality. Subjectivists hold that it is a matter of holding beliefs that promote the achievement of an epistemic goal that subjects care about achieving. Objectivists hold that it is a matter of holding beliefs that promote the achievement of an epistemic goal that is good to achieve, whether or not subjects care about achieving it. (These are not Giere's senses of "objective" and "subjective"; we are back to the sense of those terms that we have been working with all along.)

The problem of the value of the epistemic goal therefore comes in two closely related forms, one for subjectivism, and the other for objectivism. The problem for the subjectivist is that there are cases where subjects do not care about achieving the epistemic goal, either because the beliefs in question are completely useless, or because having a true belief on certain topics is positively harmful. I will proceed to talk primarily

about the subjective value of the epistemic goal. The objective value of true beliefs will come up in the objections and responses.

Examples of perfectly useless beliefs are beliefs about the number of grains of sand that will be found in the average cubic foot of sand in Waikiki Beach. Or about the total number of grains of sand on the beach. Or about a disjunction consisting of propositions about the number of grains of sand on Waikiki Beach and the precise location of the Pope in relation to the mathematical center of the solar system.<sup>59</sup> Piller gives an example of a useless truth that is hard to beat: “no one is (or ever will be) interested in the string of letters we get, when we combine the third letters of the first ten passenger’s [sic] family names who fly on FR2462 to Bydgoszcz no more than seventeen weeks after their birthday with untied shoe laces” (2009b, p.415).

Propositions like the one Piller talks about would be perfectly useless to believe: no one in fact wants to have beliefs about them, nor is it valuable to have beliefs about them. (Or: maybe *someone* would like to have beliefs about them. But I surely do not, and I imagine that I am typical in that respect.) Piller (ibid.) claims, in my view correctly, that there are beliefs of this kind that people are completely indifferent about – they are indifferent between having a true and a false belief on the topic, and they are indifferent between having no belief and having any belief at all. And yet, even in such cases, one

---

<sup>59</sup> Grimm (2008) appeals to useless beliefs like these in his argument against standard versions of the instrumental conception of epistemic rationality. Goldman (1999, p. 88-89) employs similarly useless beliefs in order to motivate the restriction on the assessment of “V-value” (veritistic value, the value of a body of beliefs relative to the truth-goal) to those propositions in which agents are interested in some way.

might acquire evidence regarding those propositions, in which case one might acquire epistemically rational beliefs about them.

Such cases spell trouble for the subjective instrumental conception, because they show that there are propositions the truth of which most (if not all) agents do not care about, so it is not the case that agents in fact have the epistemic goal, in the sense of wanting to achieve it. The subjective instrumentalist holds that having an epistemically rational belief consists of having a belief that appropriately promotes the achievement of an epistemic goal that the agent cares about achieving. If there is no such goal that subjects have with respect to some proposition, then the subjective instrumentalist has to say that that belief cannot be epistemically rational for those subjects. But if those subjects should happen to acquire evidence about the propositions in question, then they *can* form epistemically rational beliefs about them, and the subjectivist instrumentalist cannot accommodate that.

In the face of this objection – that there are truths that people just do not care about – the subjective instrumentalist might try to retreat to a counterfactual analysis. The idea would be that, *if* someone cared about these useless truths, *then* it would be instrumentally rational for them to believe them. But recall the argument that I gave against the hypothetical instrumental conception of epistemic rationality in Chapter 1: some beliefs are epistemically rational, and some epistemically irrational, in the actual world. And some such beliefs are in truths that people in the actual world have no interest in having. A counterfactual instrumental analysis of epistemic rationality can only say

that in a world where the subject in question cares about the truth, the belief will be rational if it fares well vis-à-vis the truth-goal. It cannot say that the belief is instrumentally rational in the actual world.

Another way to say that is to point out that in order for some action or belief to be instrumentally rational, because it promotes the achievement of some goal, the goal in question must be valuable in some sense. I am open to different ways that a goal might be made valuable – it can be valuable as a means to achieve a further valuable goal, or it can be valuable for its own sake, or it can be valuable just because someone desires it. If a goal is valuable in any of these ways, then it can make taking the means to achieve it instrumentally rational. But in the kinds of cases we are considering, it seems that the goal is not valuable (it is not desired, nor is it good to achieve for its own sake<sup>60</sup> or for the sake of something else that is itself valuable), and so it is not capable of rendering the means to achieve it instrumentally rational. Granted, the means *would* be instrumentally rational with respect to the epistemic goal, *if* it were valuable in some sense to achieve in these cases, but the counterfactual value of the epistemic goal is not enough to generate actual instrumental rationality. Consider an analogy with the goal of getting myself a glass of orange juice: as a matter of fact, I do not feel like having orange juice right now; I've already had too much today. But *if* I wanted to have a glass of orange juice, *then* the fact that I remain sitting here now would be instrumentally irrational with respect to that goal. Still, my remaining sitting here is not instrumentally irrational in any sense – or at

---

<sup>60</sup> We'll get to a discussion of the objective value of truth shortly – for now, I only want to register the claim that *I* at least see no objective value inherent in these truths.

least the goal of getting a glass of orange juice does not make it so – because I do *not* want to get one. That goal is not valuable to achieve right now, so it does not make my action instrumentally irrational because my action fails to promote its achievement.

Before considering objections to the argument that the epistemic goal is not valuable in the cases we are considering (we will see a number of objections), let us go through two other cases that have been influential in recent epistemology.

### **The Movies**

Thomas Kelly's (2003) movie-going case goes roughly as follows. Suppose that you want to see a movie that has been in theatres for a little while. Your friends have seen it, but you try your best to avoid having them spoil the ending for you. Nevertheless, one of your friends inconsiderately blurts out the ending in your presence: the butler did it. Now you believe that the butler did it, and your belief is epistemically rational, as it is supported by a good epistemic reason: your friend has seen it and has said that the movie ends that way.

You do not, however, *want* to have a true belief in the case at hand, because knowing how the movie ends ahead of time interferes with the enjoyment of watching the movie. Nor is having a true belief about the end of the movie valuable itself; indeed, it is positively *disvaluable*, insofar as it will interfere with an episode of harmless pleasure achievement (if we assume that achieving pleasure in harmless ways is valuable). This case is more serious than the cases of merely useless true beliefs, because it involves a belief that makes things positively *worse* for the subject.

Finally, recall the case of Alexei and Anna, which we saw in the introductory chapter:

### **The Unfaithful Wife**

John Heil's (1992) example of the unfaithful wife is a stock example of an epistemically praiseworthy but practically disastrous true belief. Although we have already seen the case in the introductory chapter, let's go through it again here. Alexei loves his wife Anna, and values his marriage as highly as he values anything. He would be destroyed if he was to get divorced. Now, he comes across very good evidence that indicates that Anna is cheating on him. Alexei knows himself and his dispositions, and he knows that if he does not believe that Anna is faithful to him, he will not be able to prevent himself from behaving coldly towards her, which would precipitate a divorce. So the question is: if Alexei forms the belief that Anna is unfaithful, is his belief epistemically rational? Or if he continues to believe that she is faithful to him, is his belief epistemically irrational?

This case is like the movie-going case, except that the consequences of having the true belief that Anna is cheating on him will be very serious for Alexei. Alexei does not want to have the belief in question, even if it is true; he would prefer to keep his false belief. He therefore does not want to achieve an epistemic goal here. Nor is achieving the epistemic goal valuable: it would be disastrous for Alexei to have a true belief here. Nevertheless, if he does form the belief that Anna is cheating on him, that belief will be well-supported by the evidence, and it will therefore be epistemically rational. Or if he

persists in his belief that Anna is faithful, that belief will be epistemically irrational, in spite of the fact that Alexei does not want to achieve the epistemic goal here.

## **7.2. Some objections**

The line of argument, then, is that there are cases of epistemically rational beliefs that are either useless or else positively bad to have even if they are true; but the instrumental conception requires that the beliefs must be appropriately related to some epistemic goal (either one that the agent wants to achieve, or else one that is valuable to achieve whether or not the agent wants to achieve it) in order to count as epistemically rational or irrational; so the instrumental conception cannot account for the epistemic rationality of the beliefs in those cases. There are a number of objections to that line of argument that I want to discuss.

### **7.2.1. Conflicting Goals, Objective Value**

One objection to consider is that it is not the case that the agent *lacks* the epistemic goal in the cases at hand, only that there are conflicting and overriding practical goals.

This objection requires two responses: one for the subjective instrumental conception, and one for the objective. Regarding the subjective conception: it seems dogmatic and a little bit desperate to claim that everyone really does value having any given true belief, at least just a little bit, as long as it is true. I am inclined to side with Kornblith when he writes that “any attempt to gain universal applicability [of epistemic norms] by appeal to goals that all humans in fact have will almost certainly run afoul of the facts. Humans beings are a very diverse lot; some of us are quite strange” (2002,

p.150). The things that people in fact want to achieve are many and varied, and it is unlikely that we all value having any given belief, at least a little bit, just as long as it is true. Perhaps I can only speak for myself, but there is an infinite number of true propositions about which I am not the least bit interested.

Some instrumentalists might want to say that we really do all want to have any given belief if it is true – it is just that that desire is extremely small in the case of relatively uninteresting beliefs, and is easily overridden, given the limits on our time and memory. But it is not difficult to construct a case where time and memory limits fall out of consideration. Sosa (2002), for example, describes a case where you have to sit in a dentist's office, and someone has stolen all the magazines. You are not sleepy. There is a phone book ready to hand; you could begin stocking your mind with many new true beliefs. Presumably, if you value all true beliefs just insofar as they are true beliefs, then you would have some inclination to reach for the phone book. But you do not do it – you'd rather just sit there and do nothing. This looks like a case where other competing interests are completely absent; you would simply rather do nothing than learn random phone numbers. (I know I would.)

In the face of cases like this, where people seem simply not to care about having a given true belief, proponents of the objective instrumental conception might want to say that they can handle such cases, because they hold that the value of the epistemic goal is independent of people's actual desires.

There are three ways that I am aware of to try to argue for the objective value of all true beliefs. I will explain and respond to each in turn.

*i. Instrumental value of true beliefs*

A common way to argue for the objective value of all true beliefs is to point out that any given true belief can become useful for achieving our practical goals. True beliefs are all valuable, on this view, because they all stand in some possible instrumental relation to the achievement of goals that we do value. So, even if we do not in fact want to achieve the epistemic goal, we really should want it, or at least, it would be good for us to achieve it. Any given true belief could become useful.<sup>61,62</sup>

This is fair enough: given the right circumstances, any given true belief could become useful. But two comments are in order here. First, there are true beliefs that are instrumentally disvaluable, as in Kelly's and Heil's cases. Appealing to the value that true beliefs have as means for achieving our other goals fails to generate any value for true beliefs that get in the way of the achievement of our goals. And second, even though any given true belief could become useful, the claim that the epistemic goal is always valuable to achieve does not follow from that. For one thing, any given *false* belief could be useful as well. (I falsely believe that a meteor will hit me if I do not move *right now*; I move; an assassin's shot misses my head). The argument that the right circumstances can

---

<sup>61</sup> Cf. Foley (1993, p. 17).

<sup>62</sup> Although this is an instrumental justification of the value of true beliefs, it is still an objectivist style of justification, because it makes the value of true beliefs independent of people's actual desires. Note, in particular, that a true belief can have instrumental value in the service of the goals that people actually desire, or those that are good for people to achieve independent of their actual desires. Such instrumental value can be objective, in the sense that even if someone does not desire to achieve it, it can still be valuable for achieving other goals that she does want to achieve.

make any true belief useful proves too much, because it trades on the action-guiding aspect of belief, rather than on the truth of belief. The argument does not provide any reason for thinking that *true* beliefs are all at least a little bit useful; it provides reason for thinking that *any* belief might become useful. And that goes nowhere at all as far as giving us reason to think that the epistemic goal is valuable to achieve, because the epistemic goal essentially involves avoidance of false beliefs.

*ii. Mill and intrinsic value*

Avoiding the instrumental justification of the value of true beliefs, then, we might try to find an argument for thinking that true beliefs are valuable in themselves. Claims of intrinsic value are notoriously difficult to justify or refute – when we cannot offer a justification of the value of some possible good in terms of how it helps to achieve some further good, then we must either just take the possible good to be a real good in itself, or else we take it not to be good. But how can we decide between these options?

One suggestion about how to proceed comes from John Stuart Mill:<sup>63</sup> look and see whether the large majority of people think that the thing in question is intrinsically valuable, or at least whether the large majority of those who are educated about it think that the thing in question is intrinsically valuable. If we find widespread agreement that the thing in question *is* valuable for its own sake, then we could take that as evidence for the claim that it is valuable for its own sake. So, in the case at hand, we should look and see whether those who are educated about it – epistemologists – take truth to be valuable

---

<sup>63</sup> “the sole evidence it is possible to produce that anything is desirable [as an ultimate end], is that people do actually desire it” (1863/1996, p. 197).

in itself. But there are highly educated epistemologists who deny that all truths are valuable to have just insofar as they are truths – for example, Goldman (1999) (in his (1986) Goldman claims that we do value truths just out of curiosity – but that’s not to say that we value *all* truths just out of curiosity); Sosa (2002) (Sosa does allow that all true beliefs may have *vanishingly* small value, for the sake of argument, but that does not appear to be his own view); Kelly (2003; 2007); Piller (2009a; 2009b); Grimm (2008). So it is not clear that all or even most people desire all true beliefs for their own sake, and so we cannot offer widespread agreement about the intrinsic value of truth as evidence of its intrinsic value.

*iii. Lynch and intrinsic value*

Finally, there is an interesting argument for the intrinsic value of true beliefs due to Michael Lynch. For Lynch, caring for truth in general is a necessary condition of having intellectual integrity, and intellectual integrity is partly constitutive of a happy or flourishing life. Caring for the truth in general is therefore partly constitutive of a good or happy life. A happy life is valuable for its own sake, so what is constitutive of a happy life is valuable for its own sake (see Lynch 2004, ch.8, esp. p.136).

A few words to explain the argument are in order, before offering any objections. First of all, for Lynch, intellectual integrity involves caring about the truth for its own sake. That means (1) not taking a stand in defense of whatever views one just happens to come across; (2) being willing to pursue the truth, i.e. not resting content with common or fashionable opinions; (3) standing for what one believes precisely because one thinks it is

true (as opposed to, say, legal); (4) being open to the truth, in the sense that one is willing to admit when one is wrong (2004, pp.132-134).

The heart of the argument for the value of truth for its own sake is that “to have intellectual integrity is to be willing to stand for one’s own best judgment on *any* matter of importance – not just when it is convenient to do so. And since, as far as we know, almost any matter could be important at some point or other, to have integrity means caring about the truth in general” (ibid., p.133).

I propose to set aside the questions of whether intellectual integrity is necessary for a happy life, or just how we should construe happiness; those thorny questions would take us too far afield. Without addressing those questions, though, there is still an important objection to register here. The objection is that it is not necessary for having intellectual integrity that one care about the truth in general.

The argument for thinking that intellectual integrity *does* require caring about the truth in general depends on the claims that as far as we know, almost any matter could become important. But that claim is not quite right. If we want to allow for the possibility of fallible knowledge – knowledge the probability of which is below absolute certainty – then we *can* know, about many apparently unimportant propositions, that they will never become important. In Kelly’s movie-going case, as well as in Heil’s case of Alexei and Anna, it is no doubt plausible to think that the truth in these cases might indeed become important – in the movie-going case, perhaps a trivia question will come up about how the movie ends; in the case of Alexei, it is plausible to say that the situation is already an

important one, given that marriages seem like important things, at least to those who are in them. But in Piller's case, for example – about “the string of letters we get, when we combine the third letters of the first ten passenger's family names who fly on FR2462 to Bydgoszcz no more than seventeen weeks after their birthday with untied shoe laces” (2009b, p.415) – we can know that this truth will never be important. This is because the likelihood that it will ever become important is so extremely low (and we know that it is so low). It is difficult to imagine any even remotely plausible scenario in which that truth might become important.

So let us take Piller's example as a good case where we know that a proposition (whether true or false) will never be important. A person might have intellectual integrity without caring about that proposition at all: one might be the kind of person who (1) does not take a stand on whatever views she happens to come across; (2) does not rest content with popular or fashionable opinions; (3) stands for what she believes in, precisely because she believes it to be true; and (4) is willing to admit when she is wrong – and yet not care about the useless truth that we are considering. Having intellectual integrity, then, does not require caring for *any* given truth, just so long as it is true.

What is compelling about Lynch's view is that it really does seem that intellectual integrity requires being willing to stand for what one believes in, just because one believes that it is the truth, in cases that are important. But the important point here is that many truths are not important, and we can know them to be so – and therefore, the

inference from caring about truth when the truth is important, to caring about the truth in general (i.e. to caring about any given true belief) just does not follow.

### **7.2.2. Some persistent criticisms**

“Oh come now,” someone might insist, “you just cannot deny that *any* true belief *can* be useful, given the right circumstances. So we really *ought* to value having any true belief whatsoever, even if we do not do so, and we *would* value all true beliefs if only we thought about it.”

I am happy to agree that given the right circumstances, any given true belief *could* become useful. But the claim that any given true belief is in fact valuable to have, or that true beliefs are always better than false ones, does not follow from that. As I have said, any given *false* belief could be useful as well. The argument that the right circumstances can make any true belief useful proves too much, because it trades on the action-guiding aspect of belief, rather than on the truth of belief. The argument does not provide any reason for thinking that *true* beliefs are all at least a little bit useful; it provides reason for thinking that *any* belief might become useful.

My imaginary critic continues to press me: “Ok, but in cases where false beliefs are useful, that’s only because you lack other true beliefs. If you had *all* of the true beliefs that there are to have, then false beliefs would no longer be useful for anything.”

I have two things to say here. (1) It is unlikely that you can have all of the true beliefs that there are to have – not even if, per impossibile, you have infinite time and infinite mental capacities. The reason is that there are uncountably infinitely many truths

to be had. The set of real numbers, for example, is an uncountably infinite set (where a countably infinite set is one that can be put in a one-to-one correspondence with the natural numbers, and an uncountably infinite set cannot). Because there is an uncountably infinite set of real numbers, it follows that there are uncountably infinite numbers of truths about the real numbers (given that there are propositions that are uniquely true of each number). Even if we could learn one thing at a time, for an infinite amount of time, and remember all of it, we still just couldn't get all the truths that there are.

(2) The second thing to say is that having all of the true beliefs that there are to have would not obviously be a good thing. For example, if we all knew all of the unsavory details about the way people think and act in private, or all of the truths about the way that human bodies function, functional human interaction might very easily break down. People might cease to be attracted to each other, or cease to have any self-respect. If that were the case, then we'd be better off not knowing all truths. (Of course that is all speculative, but I permit myself the speculation, given that it is in response to the claim that it would be practically useful to have all of the true beliefs that there are to have, which is itself an extremely speculative claim.)

“But still, it would be *epistemically* good to have all of the true beliefs. You've only argued that it's not *practically* good.” Well, ok. I was indeed concerned here with whether all truths are at least somewhat practically useful. Some instrumentalists appeal to that in arguing for the value of the epistemic goal. On the other hand, some epistemologists simply assume that all truths have a kind of value just insofar as they are

true, and they call the category of value in question epistemic value. But in the face of the kind of uninteresting truths that we have been looking at, and in the absence of positive arguments for thinking so, it is dogmatic to say that there *just is* value in beliefs just insofar as they are true. (I do not say that it is *question-begging* to assert this, because the kind of claim I am talking about is not supported by arguments – not even question-begging ones.)

Positive arguments for the value of true beliefs just insofar as they are true, which do not appeal to the practical usefulness of true beliefs, are hard to come by. However, we do find some such arguments in Lynch's work. I considered what I take to be the best one: the argument that caring for truth in general is a necessary component of a happy life, and I argued that that argument is lacking.<sup>64</sup>

All of that is to say that, if the goal is to believe *all* and *only* truths, right now, then no one in fact wants to achieve that goal. When there are overriding reasons to not achieve it, or when the beliefs in question are just too trivial, people do not care about achieving the epistemic goal. Nor is achieving the epistemic goal always valuable independently of what people desire.

---

<sup>64</sup> Another argument that Lynch gives is that caring for the truth is essential for a thriving liberal democracy (2004, ch.10). I do not propose to consider that argument here, as it has more to do with institutional norms and policies than it does with whether all truths are valuable in themselves – and also, because I would give much the same reply to that argument as I did to the argument from happiness: caring for the truth in the sense required for happiness/liberal democracy does not entail caring for *any* truth, just insofar as it is true; it entails only caring for any truth that is or might reasonably be expected to become important. Furthermore, people who do not live in liberal democratic societies can of course have epistemically rational beliefs, without having a reason to care for truth based on its value for liberal democratic society.

“Permit me one more objection: admittedly, no one (or at least, very few people) really cares about having *any* given true belief, just insofar as it is true. Still, we can *idealize* by abstracting away from the particular truths that people care about and goals that they want to achieve, and do our epistemic theorizing in a general way with reference to the epistemic goal as we have stated it. In that case, the particular epistemic goals that people really do care about will be what determine the epistemic rationality or irrationality of their beliefs, but our formulation of the epistemic goal can orient our epistemic theories.”

My response to this objection is threefold: (1) there are an infinite number of truths that people are not interested in, so to set aside truths that people are not interested in when we do our theorizing is to do more than just a little abstraction. It is easy to show that there are an infinite number of uninteresting truths. If we accept that Piller’s example is a case of an uninteresting truth, then we can provide a recipe for cooking up uninteresting truths that can be repeated ad infinitum. Recall Piller’s example: “no one is (or ever will be) interested in the string of letters we get, when we combine the third letters of the first ten passenger’s family names who fly on FR2462 to Bydgoszcz no more than seventeen weeks after their birthday with untied shoe laces” (2009b, p.415). We can generate new uninteresting truths by simply adding letters to the nonsense string of letters that Piller uses in the example. (It is easy enough to come up with other recipes for constructing infinite numbers of uninteresting truths, but we need not dwell on them any more – since, after all, they are very uninteresting.)

(2) The second response to this objection is that if we do our epistemological theorizing by abstracting away from the goals that people in fact want to achieve, or that are in fact valuable to achieve, then our theories will not be adequate to capture the epistemic rationality of people's actual beliefs. This is important: we are trying to give an account of the epistemic rationality of beliefs (and other doxastic states). If the goals that make people's beliefs epistemically rational are different from those that orient our theorizing, then our theorizing will quite miss its point.

(3) Furthermore, recall (from Chapter 4) that there is a reason why we formulated the epistemic goal in the way that we did: if it is not formulated in this way, then we will be able to come up with cases where good evidence fails to generate epistemic rationality, and where belief held against good evidence will get to count as epistemically rational. We wanted to avoid that result.

### **7.2.3. Sosa and the final value of a domain of evaluation**

A suggestion that deserves a hearing at this point is Sosa's (2007) notion that there are independent critical domains of evaluation, which have their own fundamental values which orient the evaluations that we make inside the domain – but these fundamental values need not themselves have any value outside of their domains. For example, take the case of the critical domain of evaluation associated with making good coffee. Good coffee is the fundamental value of the domain, and evaluations in the domain of coffee-making are made by reference to the way that they promote the final value of producing good coffee. And such evaluations do not themselves presuppose that good coffee is itself

valuable. Or consider another case. An expert on guns who undergoes a fundamental conversion, and sees no more value in the goal of promoting good gunmanship, can still make instrumental evaluations in the critical domain of evaluation oriented around the goal of good gunmanship.

There are two problems with this way of approaching epistemic evaluations. First, Grimm (2009) argues, it fails to capture the special importance of truth as the goal in light of which epistemic evaluations (according to an instrumentalist) are to be made. If there is no need to take truth to be itself valuable in any sense, in order for it to ground instrumental epistemic evaluations, then there is nothing to mark it off from any other goal in terms of which we might evaluate truths. All we can say is that truth is of fundamental *epistemic* value. But what then do we mean, except that truth is of value in the domain that makes its evaluations in light of the goal of achieving true beliefs? We could make the same move with respect to any other goal and domain of evaluation. (What if we just go ahead and call another domain epistemic?) Why do truth and its associated domain of evaluation have a privileged status for the evaluation of beliefs, over other goals? So the first problem here is that this approach to epistemic value fails to make room for the special importance that instrumentalists want to place on the epistemic goal as a goal in terms of which we evaluate beliefs.

There is a second problem that I want to dwell on briefly. Sosa's approach to epistemic value and epistemic evaluations is much like the hypothetical instrumental conception of epistemic rationality that I argued against in Chapter 1. It holds that we

make instrumental evaluations in light of a goal that is not itself valuable in any sense (it is neither desired nor good independent of what people desire). But that is not a plausible way to make evaluations of instrumental rationality.

Consider again the analogy with the domain of evaluation that centers on good coffee as a final value. Even granting that coffee is not itself valuable,<sup>65</sup> we can make evaluations of how effective certain actions, machines, institutions, attitudes, etc., are in light of good coffee taken as a final value. But this does not show that such actions, machines, etc., are instrumentally *rational* insofar as they promote the goal of making good coffee – for they might be carried out by people who have no interest in producing a good coffee. Someone who hates coffee might, for example, brew a coffee, with the goal in mind of producing a dark liquid that he can pretend is motor oil. His action would not be appropriately evaluated as rational or irrational with respect to the goal of producing good coffee – for that is not his goal. Perhaps even more importantly, this approach to domains of evaluation does not give anyone any kind of *reason* to take the means to produce good coffee, if they do not have a desire to make a good coffee. Similarly, in the case of the converted expert on gunmanship, the goal of promoting good gunmanship can give the expert no reason to take any means to promote that goal, for he sees no value in it (supposing also that there is no value in good gunmanship independent of people's desires).

---

<sup>65</sup> Perhaps I might grant that coffee is not valuable, in some distant possible world...

So too, in the case of epistemic evaluation, if we take truth to be the orienting value for epistemic evaluation, without taking truth itself to be valuable in some important sense, the goal of achieving true beliefs can give no one a reason to do anything, such as to form beliefs that they take to achieve that goal. The epistemic goal can only give anyone such reasons if it is valuable to achieve that goal in some sense.<sup>66</sup>

#### **7.2.4. White and hypothetical rationality**

Roger White (2007) objects to Thomas Kelly's argument against instrumentalism. White grants that instrumentalists in Kelly's sense cannot account for the epistemic rationality of the subject's belief in the movie-going case, but he holds that instrumentalists can safely retreat to the hypothetical sense of instrumental rationality: *if* the agent in question were to have the epistemic goal – *if* she wanted to believe the truth in the case at hand – *then* she would have epistemic reason to form the belief.

I have already argued that this sense of instrumental rationality is implausible. I only mention White here because he offers another analogy in support of the view, which is worth considering. The analogy is with financial rationality: not everyone always wants to be financially rational, but for someone who does not care about being financially wise, and who is content to coast by without making any investments or putting away savings, we can still say that she *would* be financially rational, *if* she was to start putting some

---

<sup>66</sup> I think a natural move to make at this point would be to say that talk of epistemic value is just a *façon de parler*, not to be taken literally in every case. Perhaps that is a fair point. But if that is the case, then epistemic value (and the epistemic goal, which is a subset of the epistemic values) cannot really ground epistemic rationality – unless, of course, talking in terms of epistemic rationality is itself just a *façon de parler*. But I am assuming that many beliefs really are epistemically rational or irrational, in the actual world.

money into savings, or to go ahead and invest. She would be financially rational, because if she cared about making good financial decisions, then those would be the things to do. That is the sense in which an agent's financial decisions can be financially rational, even if she does not care to achieve any financial goals. So we can make judgments of rationality even in spite of an agent's goals.

The appearance of plausibility that attaches to the sense of financial rationality, I would suggest, is only there because we take money to have objective instrumental value (at least to a certain extent). Without money, one cannot buy food, pay for a place to live, buy clothing, etc. Having enough money helps people to live well, and there is (intrinsic) objective value which attaches to living well. So we can say that there is a sense in which decisions that promote financial ends are instrumentally rational, in the service of a goal that is good to achieve, whether or not people care to achieve it. Such evaluations make sense because of the objective instrumental value of money. But nothing follows about the plausibility of judgments of instrumental rationality when the goal to be promoted is not valuable in any sense.

#### **7.2.5. Kornblith and desire in general**

Hilary Kornblith (2002) holds that it is highly unlikely that there will be a goal that all agents desire to achieve that can ground epistemic rationality. People are just too diverse with respect to the things that they desire. Nevertheless, Kornblith grounds epistemic norms in desire – not in any specific type of desire, but rather in the fact that agents desire anything at all. Insofar as agents intrinsically desire anything (i.e. desire anything for its

own sake), they are committed to trying to get the truth; having true rather than false beliefs will better allow us to achieve those things that we do value.<sup>67</sup> “This provides us with a reason to care about the truth whatever we may otherwise care about. It also provides us with a reason to evaluate our cognitive systems by their conduciveness to truth” (2002, p.158).

There is something true here and something misleading. What is true is that if we care about anything at all, then we must be committed to having at least some true beliefs about what we care about. Another truth here is that cognitive systems that are conducive to truth will do better at getting us those truths that are important to us. But we cannot infer that we therefore have an interest in acquiring beliefs in *any* given truth; there are truths that are entirely uninteresting to us, and that have no practical bearing on anything that we intrinsically care about. (In fairness to Kornblith, I do not see him explicitly making that inference, but it is a natural one for a defender of instrumentalism to want to make.) The point, then, is that we are not committed to trying to believe the truth with respect to topics that are entirely uninteresting to us, simply in virtue of the fact that we are committed to trying to believe the truth (or at least some truths) regarding those subjects that we do care about.

#### **7.2.6. Steglich-Petersen and teleological reasons**

Finally, let us consider Asbjørn Steglich-Petersen’s (2011) novel response to the value-based objection to the instrumental conception of epistemic rationality. He admits the

---

<sup>67</sup> Zagzebski (2009) makes a similar move, though she is explicit about not trying to ground the normativity of a universal truth-goal; she is only interested in the claim that we are committed to wanting the truth regarding those matters in which we are interested and in which we morally ought to be interested.

force of the objections that we have been considering: there are cases where agents have good epistemic reasons that support their beliefs, even though they do not *want* to have a true belief, and even though there is no objective value attaching to having such true beliefs.

Steglich-Petersen responds on behalf of the teleological approach to epistemic rationality by distinguishing two kinds of reasons for belief: (1) reasons to form a belief *about* a proposition, and (2) reasons to think *that* a proposition is true. The first type of reason is practical; the second, epistemic. Steglich-Petersen is content to allow that epistemic reasons might be categorical; paradigm cases of reasons to think that a proposition is true involve having evidence for it, and the evidential relation does not look like an instrumental one. But epistemic reasons only *give* us reasons to form beliefs (i.e. they are only normative) in the context of an all-things-considered reason to form a belief *about* a proposition, and that kind of reason is teleological, because it involves an appeal to the *value* of having a belief about the proposition. So the proposal is that if an agent has an all-things-considered-reason to form a belief about a proposition, then the agent has a reason to pursue the epistemic goal and form a true belief.

This response does handle some of the problems that we have been considering. In particular, it handles Kelly's movie-going case nicely, because the agent in that case has no all-things-considered reason to form a belief *about* the proposition that the movie will end as his friend says that it will; he has reason *not* to form such a belief. Nevertheless, he has epistemic reasons that support his belief when he does form it, and so his belief is

epistemically rational. This response also handles the problem of worthless true beliefs, because there is no all-things-considered reason to form beliefs about extremely trivial propositions, even though there can be evidence, and therefore epistemic reason, to think that such propositions are true.

The problem with Steglich-Petersen's distinction between reasons to form beliefs about a proposition and reasons to think that a proposition is true is that on this account, epistemic reasons – reasons to think that a proposition is true – are quite explicitly not themselves value-based:

Whether I have epistemic reason to believe a particular proposition depends on the epistemic properties of that (potential) belief only, and these properties are entirely unaffected by the interest I may or may not have in that proposition, the moral or instrumental worth of believing the proposition, or any other such value giving property. (2011, p.22)

Whether someone has epistemic reason to believe a particular proposition depends on the evidence available concerning the truth of that proposition. (ibid., p.30)

This is not exactly a problem for the distinction itself, but it means that the distinction between reasons to form a belief about a proposition and reasons to think that a proposition is true does not help out the instrumental conception of epistemic rationality. A person can have epistemic reasons, on this view, whether or not the epistemic properties of the case are valuable in any sense. But that is just to say that epistemic reasons are not instrumental – they do not have to do with the promotion of some sort of epistemic value. The instrumental aspect of reasons for belief, on Steglich-Petersen's account, is in the all-things-considered practical consideration about whether a person has a reason to form a belief *about* a proposition. (So his paper is perhaps not appropriately

titled: it is a defense of a teleological view of all-things-considered practical reasons for belief, but not of epistemic reasons.)

## **Chapter 8: A Regress Argument Against the Instrumental Conception**

In this chapter, I elaborate and defend an objection against instrumentalism, due to Siegel (1996a, 1996b).<sup>68</sup> This is the final argument that we will consider against instrumentalism.

### **8.1. A Restriction on Instrumental Rationality**

Siegel's question is the following: given means M, evidence E, and goal G, how is it that M can be instrumentally rational as a means to achieve G? The answer is that E must make the following claim rational to believe: "M is an effective means to achieve G" (call this claim 'C'). If E does not make C rational to believe, then M is not rational to adopt as means to achieve G. The mere fact that M will achieve one's goals is not enough to render the adoption of M rational; it must also be rational for one to think that M will do so. The point is perfectly general: for any means M (whether it be a belief or an action) and goal G (be it a practical or an epistemic goal), it cannot be instrumentally rational to adopt M in order to achieve G unless the available evidence makes it rational for the agent in question to think that M will achieve G.

To illustrate the restriction, consider an example.

#### **Pizza Meister**

Jim very badly wants to order a pizza. He knows that the nearby pizzeria, Pizza Meister, has been closed for renovations for the past week, and that there is a sign on the building, saying that the closure will last for two weeks. However,

---

<sup>68</sup> I take some liberties in the exposition of Siegel's argument, but the broad lines are the same. (In particular, Siegel does not cast the argument explicitly as a regress, except in a passing footnote (1996a, p. S123n).)

unbeknownst to Jim, the renovations ended surprisingly early, and Pizza Meister is now open, though they have not yet advertised their re-opening. There is another pizzeria that Jim likes almost as well, Chez Pizza, which he saw was open as he drove past on his way home. Jim has only one phone call left before his telephone rates skyrocket.

So: here we have a goal, to get a pizza. We have two mutually exclusive possible ways of trying to achieve that goal: to call Pizza Meister, or to call Chez Pizza. We have some available evidence, which is that Pizza Meister has been closed for one week, that they have said that they would be closed for two weeks altogether, and that Chez Pizza claims to be open. The instrumentally rational thing to do here is to call Chez Pizza: the evidence indicates that doing so will achieve the goal of getting a pizza, and it also indicates that the alternative course of action, calling Pizza Meister, will not achieve the goal. Because the evidence indicates to Jim that calling Pizza Meister will fail to achieve his goal, it would be irrational of him to do so, despite the fact that that course of action *would* in fact achieve his goal.

This restriction on the conditions under which it can be instrumentally rational to adopt some means to achieve a goal gives rise to a regress argument against instrumental conceptions of epistemic rationality. Before getting to the regress argument, though, a few notes are in order regarding the restriction. First of all, the restriction does not require that it be rational to believe that M is the *only* means available to achieve G. It does not even require that it is rational to believe that M is the *best* available means. Perhaps those

stronger conditions are true, but they are not required for the argument at hand. All that the restriction requires is that it be rational to think that M is *an* available means that will achieve the goal.

Second, the restriction is incompatible with the two senses of instrumental rationality identified by Giere (objective and subjective). Giere's objective sense of instrumental rationality is the sense that an action will *in fact* achieve a goal; his subjective sense is the sense that an action is *believed* by the subject to be such that it will achieve the goal. Both of these senses violate the restriction. It is not sufficient for a means M to be instrumentally rational, relative to some goal G, that M will in fact achieve G. If it is not rational to believe that M will achieve G, then taking means M in order to achieve G is not rational. Similarly, it is not sufficient for taking M to be rational with respect to G, that the subject in question believes that M will achieve G. If the subject's belief is irrational, then it is not rational for the subject to take M, since actions that are based on irrational beliefs are not themselves rational. (There is perhaps something akin to rationality at work in this case, because we are rationally required to keep our beliefs and our actions in line with each other. But we are also rationally required to have rational beliefs, and when our beliefs are not rational, our actions that are based on them fail to be rational as well.)

Third, the restriction is neutral regarding the recent debate over wide and narrow scope views of instrumental rationality. Briefly, the distinction is as follows. Take some subject S, some goal G, and some means M. If S has goal G, and M is a necessary means

to take to achieve G, then narrow-scopers hold that S ought to take means M. Wide-scopers object that the narrow-scope view entails that an agent who has immoral goals ought to take the means to achieve her goals, but it is not the case that agents ought to take the means to achieve immoral goals.

The two views can be stated as follows:

**Wide-scope**

Ought: if you have a goal, then take the means to achieve it.<sup>69</sup>

(Equivalently: either you ought to give up your goal, or else you ought to take the means to achieve it.)

**Narrow-scope**

If you have a goal, then you ought to take the means to achieve it.

(Equivalently: either you do not have a goal, or else you ought to take the means to achieve it.)

The disagreement is over the scope of the “ought.” Wide-scopers hold that it governs the whole conditional; narrow-scopers hold that it only governs the consequent. Many theorists take it as unproblematic that wide-scoping is the way to go.<sup>70</sup> However that debate settles, the point here is that, although Siegel’s restriction looks somewhat like the narrow-scope view of instrumental rationality, it is in fact independent of the wide/narrow-scope distinction. Again, the restriction is that, if you have a goal, then in

---

<sup>69</sup> It might help to read that statement as “you ought to be such that if you have a goal, then you take the means to achieve it.”

<sup>70</sup> Broome (1999; 2007) champions the wide-scope view. Schroeder (2004) argues against the wide-scope view, although he does not exactly defend the narrow-scope view, either. Kolodny (2005) defends the narrow-scope view with respect to at least some cases.

order for taking the means to achieve it to be instrumentally rational, the evidence available must make it rational to believe that taking those means will help achieve the goal. That restriction can easily be built into both the narrow and the wide-scope views.

They would then look like:

**Wide-scope + restriction**

Ought: if you have a goal, then (you take some means to achieve it only if the evidence makes it rational to believe that those means will help achieve it).

**Narrow-scope + restriction**

If you have a goal, then (you ought to take some means to achieve it only if the evidence makes it rational to believe that those means will help achieve it).

So the restriction is neutral with respect to the wide/narrow-scope distinction.

## **8.2. The Regress**

The regress argument against instrumental conceptions of epistemic rationality that arises from this restriction is as follows. Suppose that all rationality is instrumental, including epistemic rationality. Once again, the restriction is that in order for it to be rational to adopt means M in order to achieve goal G, given evidence E, it must be the case that E makes claim C (the claim that M will achieve G) rational to believe. Therefore, if all epistemic rationality is instrumental, there must be some goal with respect to which claim C is itself an instrumentally rational means to adopt. Naturally, we might propose that that goal is the epistemic goal, say, of now having true beliefs and now not having false beliefs. So there is a goal, G<sub>2</sub> (i.e. the epistemic goal), with respect to which C must be

instrumentally rational. But, given the by-now-familiar restriction on instrumental rationality, the fact (if it is a fact) that believing C is effective for achieving the epistemic goal  $G_2$  is not sufficient for it to be instrumentally rational to believe C. It must also be the case that the available evidence makes a *further* claim  $C_2$  rational to believe –  $C_2$ , of course, being the claim that believing C is effective as a means for achieving  $G_2$ . But again, if all rationality is instrumental, then the rationality of believing  $C_2$  must be relative to the achievement of some goal,  $G_3$ .<sup>71</sup> And, given the restriction on instrumental rationality that we are working with, in order for believing  $C_2$  to be instrumentally rational with respect to  $G_3$ , it must be the case that the available evidence makes a further claim  $C_3$  rational to believe –  $C_3$  being the claim that  $C_2$  is effective as a means for achieving  $G_3$ . And so on. The regress arises here because the rationality of adopting some given means for achieving a goal *depends* on whether the evidence makes it rational to believe that the means will be effective ones. So, if all rationality is instrumental, then all cases of rationally adopting means to achieve goals will depend on the rationality of adopting some *further* means to achieve a goal. We have here an infinite regress, and it spells trouble for any thoroughgoing instrumental conception of rationality.

### 8.3. Three objections

There are three objections that might be offered to this line of argument. The first is that Foley's appeal to reflective stability gives us an instrumental conception of epistemic rationality that does not require it to be rational to think that the means in question are

---

<sup>71</sup> Note that  $G_3$  might be identical with  $G_2$  – they can both be the epistemic goal. The point is only that we require an infinite series of claims C,  $C_2$ ,  $C_3$ , etc., that are instrumentally rational to believe, not that we need an infinite series of goals.

good ones for achieving the goal, so his account avoids the restriction that I have argued for. The second objection is that we do not need to worry about regresses – infinitism, for example, is an account of epistemic justification that is at home in an infinite regress of justifying reasons. And the third objection is that only internalist accounts need to build in the restriction on instrumental rationality; an externalist account can reject it.

### **8.3.1. Sufficient reflection again**

The first objection is one that I have already responded to (section 6.5), but it is important to note the objection in this context as well. The objection is that Foley's sufficient-reflection conception of epistemic rationality does not appeal to any further normative conditions to generate instrumental rationality; it only appeals to the point of sufficient reflection, which is just the point of reflective stability when reflecting from the purely epistemic point of view. Foley's general conception of rationality is as follows: "a plan, decision, action, strategy, etc. is rational in sense X for an individual just in case it is *epistemically rational* for the individual to believe that the plan, decision, action, strategy, etc. will do an acceptably good job of satisfying goals of kind X" (2005, p.318). So far, Foley's account is in line with the restriction I am proposing on instrumental rationality: it must be epistemically rational to think that the means will be acceptably good ones for achieving the goal in question. But beliefs are epistemically rational for a subject, on Foley's account, just in case they are such that the subject would take them to do an acceptably good job of achieving the epistemic goal, upon sufficient reflection. The idea here, then, is that sufficient reflection can provide a way out of the regress by blocking

the requirement that it be epistemically rational to think that a belief does an acceptably good job of achieving the epistemic goal, in order for that belief to be instrumentally rational in the service of the epistemic goal.<sup>72</sup>

It is important, then, that we be able to show that there are problems for the sufficient-reflection view of epistemic rationality. The problem with the view, I have argued, is that the sufficient-reflection account does not respect the proper basing requirement on epistemically rational beliefs.

### **8.3.2. Regresses – nothing to worry about**

The second objection is that regresses are not necessarily worrisome things. Many of us take regresses, or at least regresses that occur in a dependence relation, to be serious problems. (In the case at hand, the rationality of taking some means to achieve a goal *depends* on the rationality of a belief about the effectiveness of that means for achieving that goal.) The general idea behind regress objections is that a dependence relation that does not have a stable bottom is no dependence at all.

But perhaps those inclined toward infinitism about justification (those who allow the possibility of infinite chains of justifying reasons) would not accept that reasoning. As long as, at each further stage in the regress, a rational belief in the effectiveness of the means in question to achieve the goal in question (i.e. the epistemic goal) can be had, the regress is not problematic. If no such belief can be had, then the justification of the initial

---

<sup>72</sup> In fact, Foley does think that epistemic rationality is also an instance of the general schema of rationality, so there is a surface worry about the threat of circularity in the explanation of general rationality, but the important thing is that Foley's analysis of epistemic rationality itself does not appeal to any further notions of rationality.

belief will be in danger of being undermined, but until we reach the point where we cannot find further reasons when they are called for, the regress is all right, and the initial belief can be justified (cf. Klein 2007a; 2007b).

One problem for this infinitist objection is that, once we get to a sufficiently complex stage in the regress, the belief in question will become too complex for a subject to hold in her mind. Once we get to that point, it will not be rational to believe that the means in question are good ones for achieving the goal in question (i.e. to believe that the belief in question is effective for achieving the epistemic goal). No one could even hold the proposition before their mind, much less have any notion of whether it might be true. Once we get to this sufficiently complex level, the belief is no longer rational to adopt. So, in response to the infinitist objection that the regress is not problematic as long as we can have a rational belief at every stage in the regress: we *cannot* have a rational belief at every stage in the regress, because at some point in the regress, we cannot have a belief at all.

Another problem for this infinitist response is that it is not clear what reason one could even have for thinking that the claims at very distant stages in the regress are true. If one could not have such reasons, then at those stages of the regress, it will no longer be rational to believe that the claim is true.

One might instead try a coherentist response to the infinite regress that we are concerned with here. The coherentist response to the standard regress argument in favour of foundationalism about epistemic justification is that we do not need any ultimate,

bedrock justifiers; as long as each stage in the regress is a belief that is supported by other beliefs in one's stock or web of beliefs, and there are relations of mutual support among one's beliefs, and the web of beliefs is large enough, then it is all right if the chain of justifying reasons turns back on itself in a circular fashion.

That coherentist response does not work for the regress at hand, however, because in this regress, the claim that must be employed at each stage of the regress is always a *further* claim about the rationality of the belief at the stage of the regress that came just before. Presumably, these new beliefs that must be introduced at each new stage of the regress are not already part of a subject's belief system, nor do they enjoy roundabout support from the beliefs appealed to at previous stages of the regress. The regress of reasons will not involve beliefs that occurred at previous stages in the regress, so even if circles in our reasoning are acceptable (a controversial claim, to be sure), the chain of beliefs in this case will not circle back on itself.

If these replies do not convince you that the regress is a serious one, here is another way to cast the objection that I am trying to make, which does not trade on regresses of any sort: we are trying to give an account of the concept of epistemic justification. In any good analysis, the *definiens* must not appear in the *definiendum*: what is to be defined cannot appear in the terms that define it. But if epistemic rationality is defined in terms of instrumental rationality, and instrumental rationality essentially involves a restriction about the epistemic rationality of the effectiveness of means for

achieving goals, then epistemic rationality appears in (or at least is presupposed by) the terms that are supposed to define epistemic rationality.

This way of putting the present argument against the instrumental conception of epistemic rationality avoids objections that are suggested by the formulation of the argument as a regress.<sup>73</sup> I cast the argument as a regress, because infinite regresses strike me (as they strike many philosophers) as problematic, but the argument can be made simply in terms of the conditions of good conceptual analyses.

### **8.3.3. Externalism**

There is one further objection to consider before concluding this chapter:<sup>74</sup> the restriction required on instrumental rationality in order to generate the regress is an internalist restriction. It can be rational in an *externalist* sense to adopt means to achieve goals, even if the evidence does not make it rational for the subject to believe that the means are good ones. Externalist instrumentalists can therefore ignore the restriction and the regress both. (This objection involves an appeal to externalism in epistemology, not in metaethics/action theory or semantics. The metaethical externalist position, that moral beliefs need not be intrinsically motivating and need not have reference to a person's "subjective motivational set," which we saw briefly in Chapter 3, is a different beast altogether. So too is the semantic externalist position, which is that the reference of at

---

<sup>73</sup> Cf. Fumerton's objection to Klein's infinitism (Fumerton 2001b, pp. 7-8).

<sup>74</sup> Wrenn (2004) gives a further objection to Siegel's argument, which I do not consider here. Wrenn's argument treats Siegel's argument as addressing the question of the *value* of epistemic justification or rationality, i.e. whether its value is instrumental or categorical. Siegel's main point, however, is that epistemic rationality is not *constituted* by instrumental concerns. The value of justification is a separate question.

least some of our terms – particularly natural kind terms – is fixed by factors outside of our heads.)

The reply to this objection is twofold. First, as I have already indicated, I do not think that externalist accounts of rationality and justification are promising, and I am not very much interested in them here. But second, externalists do not get to reject the restriction so easily. Internalists about justification hold that all that is relevant to the justification of beliefs is cognitively available to, or internal to the perspective of, the believing subject. Externalists hold that some or all things that are relevant to the justification of a subject's beliefs can be cognitively unavailable to her, or external to her perspective.

Externalism, so construed, does not straightforwardly exclude the restriction on instrumental rationality. The restriction, again, is that in order for adopting means M to be instrumentally rational for achieving goal G given evidence E, E must make it rational to believe claim C, which is the claim that M will be effective means for achieving G. That restriction does not require that the *fact* that E makes C rational to believe is cognitively accessible to the believing subject; it only requires that E in fact makes C (epistemically – and so, for an instrumentalist, instrumentally) rational to believe. So the externalist instrumentalist is saddled with the regress, too: if E must make C rational to believe, and all rationality is instrumental, then even if the rationality or justification of belief need not be cognitively accessible to the believer, the problem remains that we have instrumental

rationality entering into the condition upon which instrumental rationality depends. That is what generates the regress.

So perhaps an externalist will want to say that the sense in which C must be rational to believe is not itself an instrumental sense. Very good: then this externalist is not a thoroughgoing instrumentalist about epistemic rationality.

## **Chapter 9: Conclusion**

To recap: the instrumental conception of epistemic rationality is the view that the epistemic rationality of our beliefs (or more generally, our doxastic states) depends on some relevant epistemic goal or set of goals. The objective instrumental conception holds that the goals that are relevant for the determination of the epistemic status of beliefs are independent of the desires and aims that agents in fact have. The subjective instrumental conception holds that the goals that are relevant for the epistemic evaluation of beliefs are those that epistemic agents care about achieving.

The instrumental conception is of interest to a number of epistemologists today – some think that it is a promising account of epistemic rationality, while others think that it is deeply mistaken. We saw in Chapter 3 that there are a number of reasons for accepting instrumentalism. The reason that the instrumental conception is so interesting to me is that it offers the possibility of a deep explanation of the importance of evidence for epistemic evaluations. Another reason for accepting instrumentalism is that it seems to comport so well with epistemological naturalism. I argued, however, that the reasons that at first appear to offer such strong support for instrumentalism do not in fact force us to accept it; those reasons are not as strong as we might think.

In Chapter 4, I argued that the most plausible formulation of the epistemic goal is essentially Richard Foley's: now to believe those propositions that are true and now not to believe those propositions that are false. I am not deeply committed to this formulation of the epistemic goal; I put it forward as what I take to be the most plausible formulation, in

order to give the instrumental conception the best hearing possible, before going into the arguments against it.

In Chapters 5 through 8, I went through a number of arguments against the instrumental conception. Chapter 5 was concerned with the proper-function approach to epistemic rationality, which is a version of the objective instrumental conception: it holds that the epistemic goal is normative for us, and it determines the epistemic status of our beliefs, because producing true beliefs is the proper function of (at least part of) our cognitive systems, and proper functions are normative. The two main problems with this approach to epistemic rationality are (1) that proper functions are not in general normative for us, except in special cases where there is some independent reason why fulfilling a proper function is good, and (2) it is not obvious that our cognitive systems really have the right kind of causal histories to have proper functions. And if human cognitive systems do not have proper functions, then of course proper functions cannot ground the epistemic rationality of human doxastic states.

In Chapter 6, I went through the argument that we can come up with counterexamples to the instrumental conception, where a belief is held exactly as it ought to be, considered in light of the epistemic goal, but where that belief fails to be epistemically rational, because it is held against good available evidence. This objection met with some resistance, in particular by the restriction that each belief must achieve the epistemic goal in and of itself (irrespective of any of its consequences for a person's

broader body of beliefs). This restriction handles the initial objection, but it brings problems of its own.

Chapter 7 set out the argument that any plausible instrumental account of rationality, in the service of some goal, depend on that goal's being valuable in some sense, either because agents in fact care about it, or because there is some reason to care about it independent of what agents care about. There are cases, however, where the epistemic goal is not valuable to achieve in any sense, and yet there can be paradigms of epistemically rational beliefs in such cases, when a belief is based on very good evidence.

Finally Chapter 8 contained the argument that, given a plausible constraint on when taking some given means to achieve a goal can be instrumentally rational, taking epistemic rationality to be instrumental gets us into a vicious regress. If the arguments from Chapters 5 through 8 are good ones, then the instrumental conception faces problems serious enough to warrant giving it up.

I have not attempted to establish anything more than that negative thesis in this dissertation. In particular, I have not given any positive arguments in favour of an alternative to the instrumental conception, such as traditional categorical evidentialism. It will no doubt be clear by now that I find traditional evidentialism appealing, but my aim here has not been to give direct arguments for that view. Of course, a thoroughgoing case for categorical evidentialism will have to address the plausibility of other views, such as the instrumental conception. Insofar as the arguments here show that the instrumental

conception is implausible, they also provide a resource for someone who wants to argue for a traditional evidentialism.

I am not concerned with giving a positive account of rationality here. My central claim is that taking instrumental rationality to be able to generate a general conception of epistemic rationality is implausible. However, despite the fact that my thesis is entirely negative, it is still interesting, because the instrumental conception of epistemic rationality does have some appeal, and it has a number of adherents.

Even if my arguments here do not succeed, say, because there is a way to refine one of the versions of the instrumental conception in order to avoid the objections, still, both proponents and critics of instrumentalism ought to find these arguments helpful, as they point out just where the instrumental conception is most vulnerable. And anyone interested in questions about epistemic goals and values should find the arguments in Chapter 4 helpful, where I try to bring some order to that discussion.

Furthermore, as I indicated in Chapter 3, if a version of the instrumental conception survives the arguments here, then at least we might be able to appeal to instrumentalism in order to explain the importance of evidence for epistemic evaluations, which it seems to me is a very desirable feature of a theory of epistemic rationality. However, if (as it seems to me) the arguments here are good ones, then we ought to give up the instrumental conception as a general account of the epistemic rationality of human doxastic states altogether.

## References

- Almeder, Robert (1990). On Naturalizing Epistemology. *American Philosophical Quarterly*. vol. 27, no. 4, pp. 263-279.
- Alston, William (1985). Concepts of Epistemic Justification. *The Monist*. vol. 68.  
Reprinted in his *Epistemic Justification: Essays in the Theory of Knowledge*.  
Ithaca, NY: Cornell University Press, 1989.
- (1988). The Deontological Conception of Epistemic Justification. *Philosophical Perspectives*. vol. 2: Epistemology, pp. 257-299.
- (1989). Foley's Theory of Epistemic Rationality. *Philosophy and Phenomenological Research*. vol. 50, no. 1, pp. 135-147.
- (2005). *Beyond Justification: Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.
- Aristotle. *Metaphysics*. In *Aristotle's Metaphysics*. Trans. and with commentaries and glossary by Hippocrates G. Apostle. Bloomington and London: Indiana University Press, 1966.
- Atkinson, Katie, Bench-Capon, Trevor, and McBurney, Peter (2004). Justifying Practical Reasoning. *Proceedings of the Fourth International Workshop on Computational Models of Natural Argument*. Florianna Grass, Chris Reed, and Giuseppe Carenini, eds. pp. 87-90.
- (2006). Computational Representation of Practical Argument. *Synthese*. vol. 152, pp. 157-206.

- Audi, Robert (2001). *The Architecture of Reason: the Structure and Substance of Rationality*. Oxford and New York: Oxford University Press.
- Bergmann, Michael (2006a). Epistemic Circularity and Common Sense: A Reply to Reed. *Philosophy and Phenomenological Research*. vol. 73, no. 1, pp. 198-207.
- (2006b). *Justification Without Awareness: A Defense of Epistemic Externalism*. Oxford: Oxford University Press.
- Boghossian, Paul (2006). *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Clarendon Press; New York: Oxford University Press.
- BonJour, Laurence (1985). *The Structure of Empirical Knowledge*. Cambridge, Mass: Harvard University Press.
- (1994). Against Naturalized Epistemology. *Midwest Studies in Philosophy*. vol. 19, pp. 283-300.
- (2010). The Myth of Knowledge. *Philosophical Perspectives*. vol. 24: Epistemology. pp. 57-83.
- Broome, John (1999). Normative Requirements. *Ratio*. vol. 12, no. 4, pp. 398-419.
- (2007). Wide or Narrow Scope? *Mind*. vol. 116, no. 462, pp. 359-370.
- Brown, Harold (1988). Normative Epistemology and Naturalized Epistemology. *Inquiry*. vol. 31, pp. 53-78.
- Burge, Tyler (2003). Perceptual Entitlement. *Philosophy and Phenomenological Research*. vol. 67, no. 3, pp. 503-548.

- Card, Robert (2007). Inconsistency and the Theoretical Commitments of Hooker's Rule-Consequentialism. *Utilitas*. vol. 19, no. 2, pp. 243-258.
- Cruz, Joseph, and John Pollock (2004). The Chimerical Appeal of Epistemic Externalism. In Richard Schantz, ed. *The Externalist Challenge*. Berlin: Walter de Gruyter. pp. 125-141.
- Cummins, Robert (1975). Functional Analysis. *The Journal of Philosophy*. vol. 72, pp. 741-765.
- Cummins, Robert, and Martin Roth (2010). Traits Have Not Evolved to Function the Way They Do Because of a Past Advantage. In Francisco Ayala and Robert Arp, eds. *Contemporary Debates in Philosophy of Biology*. Chichester, UK: Wiley-Blackwell. pp. 72-85.
- David, Marian (2001). Truth as the Epistemic Goal. In Matthias Steup, ed. *Knowledge, Truth, and Duty: Essays on Epistemic Justification*. New York and Oxford: Oxford University Press. pp. 151-169.
- Davidson, Donald (1963). Actions, Reasons, and Causes. *Journal of Philosophy*. vol. 60, pp. 685-700.
- (1987). Knowing One's Own Mind. *Proceedings of the American Philosophical Association*. vol. 60, pp. 441-458.
- Dreier, James (2010). When Do Goals Explain the Norms that Advance Them? In Russ Shafer-Landau, ed. *Oxford Studies in Metaethics*, vol. 5. Oxford and New York: Oxford University Press, pp. 153-174.

- Engel, Pascal (2004). Truth and the Aim of Belief. In D. Gilles, ed. *Laws and Models in Science*. King's College Publications, pp. 77-97.
- (2002). *Truth*. Montreal: McGill-Queen's University Press.
- Feldman, Richard (2000). The Ethics of Belief. *Philosophy and Phenomenological Research*. vol. 60, no. 3, pp. 667-695.
- (1989). Foley's Subjective Foundationalism. *Philosophy and Phenomenological Research*. vol. 50, no. 1, pp. 149-158.
- Foley, Richard (1987). *The Theory of Epistemic Rationality*. Cambridge, Mass: Harvard University Press.
- (1993). *Working Without a Net: A Study of Egocentric Rationality*. New York and Oxford: Oxford University Press.
- (2001). *Intellectual Trust in Oneself and Others*. Cambridge: Cambridge University Press.
- (2005). Justified Belief as Responsible Belief. In Matthias Steup and Ernest Sosa, eds. *Contemporary Debates in Epistemology*. Malden, MA: Blackwell. pp. 313-326.
- (2008). An Epistemology that Matters. In P. Weithman, ed. *Liberal Faith: Essays in Honor of Philip Quinn*. Notre Dame, Indiana: University of Notre Dame Press, pp. 43-55.

- Fumerton, Richard (2001a). Epistemic Justification and Normativity. In Matthias Steup, ed. *Knowledge, Truth, and Duty: Essays on Epistemic Justification*. New York and Oxford: Oxford University Press. pp. 49-62.
- (2001b). Classical Foundationalism. In Michael DePaul, ed. *Resurrecting Old-Fashioned Foundationalism*. Lanham, MD: Rowman & Littlefield, pp. 3-20.
- Gettier, Edmund (1963). Is Justified True Belief Knowledge? *Analysis*. vol. 23, no. 6, pp. 121-123.
- Gibbard, Allan (2007). Rational Credence and the Value of Truth. In Tamar Szabo Gendler and John Hawthorne, eds. *Oxford Studies in Epistemology: Volume 2*. Oxford: Oxford University Press.
- Giere, Ronald (1989). Scientific Rationality as Instrumental Rationality. *Studies in History and Philosophy of Science*. vol. 20, no. 3, pp. 377-384.
- (2001). Critical Hypothetical Evolutionary Naturalism. In C. Heyes and L. Hull, eds. *Selection Theory and Social Construction: The Evolutionary Epistemology of Donald T. Campbell*. Albany, NY: State University of New York Press, pp. 53-70.
- (2006). Modest Evolutionary Naturalism. *Biological Theory*. vol. 1, no. 1, pp. 52-60.
- Godfrey-Smith, Peter (1994). A Modern History Theory of Functions. *Noûs*. vol. 28, no. 3, pp. 344-362.
- (1996). *Complexity and the Function of Mind in Nature*. Cambridge; Cambridge University Press.

- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, Mass.: Harvard University Press.
- (1999). *Knowledge in a Social World*. Oxford: Clarendon Press.
- Gould, S. J, and Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London B*. vol. 205, pp. 581-598.
- Graham, Peter (2011a). Does Justification Aim at Truth? *Canadian Journal of Philosophy*. vol. 41, no. 1, pp.51-72.
- (2011b). Epistemic Entitlement. *Noûs*.
- Grimm, Stephen (2008). Epistemic Goals and Epistemic Values. *Philosophy and Phenomenological Research*. vol. 77, no. 3, pp. 725-744.
- (2009). Epistemic Normativity. In A. Haddock, A. Millar, and D. Pritchard, eds. *Epistemic Value*. Oxford: Oxford University Press. pp. 243-264.
- Haack, Susan (1993). *Evidence and Inquiry*. Oxford, UK, and Cambridge, Mass: Blackwell Publishers.
- Heil, John (1983). Believing What One Ought. *The Journal of Philosophy*. vol. 80, no. 11, pp. 752-765.
- (1992). Believing Reasonably. *Noûs*. vol. 26, no. 1, pp. 47-62.
- Hess, Marcus (2010). *Is Truth the Primary Epistemic Goal?* Ontos Verlag.
- Hooker, Brad (2007). Rule-Consequentialism and Internal Consistency: A Reply to Card. *Utilitas*. vol. 19, no. 4, pp. 514-9.

- Hubin, Donald (2001). The Groundless Normativity of Instrumental Rationality. *The Journal of Philosophy*. vol. 98, no. 9, pp. 445-468.
- Huss, Brian (2009). Three Challenges (and Three Replies) to the Ethics of Belief. *Synthese*. vol. 168, pp. 249-271.
- James, William (1949). The Will to Believe. In his *Essays in Pragmatism*. New York: Hafner. pp. 88-109. Originally published in 1896.
- Jenkins, Carrie (2007). Epistemic Norms and Natural Facts. *American Philosophical Quarterly*. vol. 44, no. 3, pp. 259-272.
- Joyce, Richard (2001). *The Myth of Morality*. Cambridge: Cambridge University Press.
- Kelly, Thomas (2003). Epistemic Rationality as Instrumental Rationality: A Critique. *Philosophy and Phenomenological Research*. vol. 66, no. 3, pp. 612-640.
- (2007). Evidence and Normativity: Reply to Leite. *Philosophy and Phenomenological Research*. vol. 75, no. 2, pp. 465-474.
- Kitcher, Philip (1992). The Naturalists Return. *Philosophical Review*. vol. 101, pp. 53-114.
- Klein, Peter (2007a). Human Knowledge and the Infinite Progress of Reasoning. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. vol. 134, no. 1, pp. 1-17.
- (2007b). How to be an Infinitist about Doxastic Justification. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. vol. 134, no. 1, pp. 25-29.

- Kolodny, Niko (2005). Why Be Rational? *Mind*. vol. 114, pp. 509-563.
- Kornblith, Hilary (1994). Naturalism: Both Metaphysical and Epistemological. *Midwest Studies in Philosophy*. vol. 19, pp. 39-52.
- (2002). *Knowledge and its Place in Nature*. Oxford: Clarendon Press of Oxford University Press.
- Korscz, Keith (1997). Recent Work on the Basing Relation. *American Philosophical Quarterly*. vol. 34, no. 2, pp. 171–191.
- (2000). The Causal-Doxastic Theory of the Basing Relation. *Canadian Journal of Philosophy*. vol. 30, no. 4, pp. 525-550.
- Korsgaard, Christine (1997). The Normativity of Instrumental Reason. In Garrett Cullity and Berys Kraut, eds. *Ethics and Practical Reason*. Oxford: Oxford University Press.
- (2009). *Self-Constitution: Agency, Identity, Integrity*. Oxford and New York: Oxford University Press.
- Kvanvig, Jonathan (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- (2005). Truth is Not the Primary Epistemic Goal. In Matthias Steup and Ernest Sosa, eds. *Contemporary Debates in Epistemology*. Malden, MA: Blackwell. pp. 285-296.
- Latus, Andrew (2000). Our Epistemic Goal. *Australasian Journal of Philosophy*. vol. 78, no. 1, pp. 28-39.

- Laudan, Larry. (1990). Aim-less Epistemology? *Studies in History and Philosophy of Science*. vol. 21, pp. 315-322.
- Leite, Adam (2007). Epistemic Instrumentalism and Reasons for Belief: A Reply to Tom Kelly's "Epistemic Rationality as Instrumental Rationality: A Critique." *Philosophy and Phenomenological Research*. vol. 75, no. 2, pp. 456-464.
- Lehrer, Keith (1990). *Theory of Knowledge*. Boulder: Westview Press.
- Leplin, Jarrett (2009). *A Theory of Epistemic Justification*. Dordrecht: Springer.
- Lewens, Tim (2004). *Organisms and Artifacts: Design in Nature and Elsewhere*. Cambridge, Mass.: MIT Press.
- Lycan, William (1988). *Judgment and Justification*. Cambridge and New York: Cambridge University Press.
- Lynch, Michael (2004). *True to Life: Why truth Matters*. Cambridge, Mass: MIT Press.
- (2009). Truth, Value, and Epistemic Expressivism. *Philosophy and Phenomenological Research*. vol. 79, no. 1, pp. 76-97.
- Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. New York: Penguin Books.
- Maffie, J. (1990). Naturalism and the Normativity of Epistemology. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. vol. 59, no. 3, pp. 333-349.
- Maitzen, Stephen. (1995). Our Errant Epistemic Aim. *Philosophy and Phenomenological Research*. vol. 55, no. 4, pp. 869-876.

- Mill, John Stuart (1996). *Utilitarianism*. In his *On Liberty and Utilitarianism*. Originally Published in 1863.
- Millar, Alan (2009). How Reasons for Action Differ from Reasons for Belief. In Simon Robertson, ed. *Spheres of Reason: New Essays in the Philosophy of Normativity*. Oxford and New York: Oxford University Press.
- Millgram, Elijah (1997). *Practical Induction*. Cambridge, Mass: Harvard University Press.
- Millikan, Ruth (1984). *Language, Thought, and other Biological Categories*. Cambridge, Mass.: MIT Press.
- (1989). In Defense of Proper Functions. *Philosophy of Science*. vol. 56, no. 2, pp. 288-302.
- Mills, Eugene (1998). The Unity of Justification. *Philosophy and Phenomenological Research*. vol. 58, no. 1, pp. 27-50.
- Moser, Paul (2002). Introduction. In Paul Moser, ed. *The Oxford Handbook of Epistemology*. Oxford: Oxford University Press.
- Neander, Karen (1991a). Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science*. vol. 58, no. 2, pp. 168-184.
- (1991b). The Teleological Notion of 'Function.' *Australasian Journal of Philosophy*. vol. 69, no. 4, pp. 454-468.

- (2009). Teleological Theories of Mental Content. *The Stanford Encyclopedia of Philosophy*. Winter 2009 Edition. Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2009/entries/content-teleological/>.
- Nozick, Robert (1981). *Philosophical Explanations*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- O'Neill, Onora (1989). *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge and New York: Cambridge University Press.
- Owens, David (2003). Does Belief Have an Aim? *Philosophical Studies*. vol. 155, no. 3, pp. 283-305.
- Perlman, Mark (2010). Traits Have Evolved to Function the Way They Do Because of a Past Advantage. In Francisco Ayala and Robert Arp, eds. *Contemporary Debates in Philosophy of Biology*. Chichester, UK: Wiley-Blackwell. pp. 53-71.
- Piller, Christian (2009a). Desiring the Truth and Nothing But the Truth. *Noûs*. vol. 43, no. 2, pp. 193-213.
- (2009b). Valuing Knowledge: A Deontological Approach. *Ethical Theory and Moral Practice*. vol. 12, pp. 413-428.
- Plantinga, Alvin (1974). *The Nature of Necessity*. Oxford: Oxford University Press.
- (1993). *Warrant and Proper Function*. New York: Oxford University Press.
- Plato. *Theatetus*. Trans. M. J. Levitt, revised by M. F. Burnyeat. In Plato, *Complete Works*. J. M. Cooper and D. S. Hucheson, eds. Indianapolis and Cambridge: Hackett, 1997. pp. 157-234.

- Meno*. Trans. G. M. A. Grube. In Plato, *Complete Works*. J. M. Cooper and D. S. Huthcheson, eds. Indianapolis and Cambridge: Hackett, 1997. pp. 870-897.
- Pritchard, Duncan. (2007). Recent Work on Epistemic Value. *American Philosophical Quarterly*. vol. 44, no. 2, pp. 85-110.
- Putnam, Hilary (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.
- Quine, W. V. (1968). Epistemology Naturalized. In his *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969, pp. 69–90.
- (1986). Reply to White. In L. E. Hahn and P. A. Schilpp, eds. *The Philosophy of W.V. Quine*. LaSalle, Illinois: Open Court, pp. 663–665.
- Quinn, Warren (1993). Putting Rationality in its Place. In his *Morality and Action*. Cambridge: Cambridge University Press, pp. 228-255.
- Reynolds, Steven (1991): Knowing How to Believe with Justification. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. vol. 64, no. 3, pp. 273-292.
- Riggs, Wayne (2003). Balancing Our Epistemic Goals. *Noûs*. vol. 37, no. 2, pp. 342-352.
- (2008). Epistemic Risk and Relativism. *Acta Analytica*. vol. 23, no. 1, pp. 1-8.
- Rorty, Richard (1998). *Truth and Progress: Philosophical Papers*. Cambridge: Cambridge University Press.
- Schroeder, Mark (2004). The Scope of Instrumental Reason. *Philosophical Perspectives*. vol.18: Ethics. pp. 337-364.

- Schwitzgebel, Eric (2010). Belief. *The Stanford Encyclopedia of Philosophy*. Winter 2010 Edition. Edward N. Zalta (ed.), URL=  
<<http://plato.stanford.edu/archives/win2010/entries/belief/>>.
- Searle, John (2001). *Rationality in Action*. Cambridge, Mass: MIT Press.
- Siegel, Harvey (1989). Philosophy of Science Naturalized? Some Problems with Giere's Naturalism. *Studies in History and Philosophy of Science*. vol. 20, no. 3, pp. 365-375.
- (1990). Laudan's Normative Naturalism. *Studies in History and Philosophy of Science*. vol. 21, pp. 295-313.
- (1996a). Instrumental Rationality and Naturalized Philosophy of Science. *Philosophy of Science*. vol. 63, Supplement: Proceedings of the 1996 Biennial Meetings of the Philosophy of Science Association. Part I: Contributed Papers, pp. S116-S124.
- (1996b). Naturalism, Instrumental Rationality, and the Normativity of Epistemology. *Protosoziologie*. vol. 8/9, pp. 97-110. Reprinted in Peter, Georg, and Preyer, Gerhard, eds. *The Contextualization of Rationality: Problems, Concepts, and Theories of Rationality*. Paderborn: Mentis, 2000, pp. 95-107.
- Smith, Michael (1995). *The Moral Problem*. Oxford, UK; Cambridge, Mass: Blackwell.
- Sosa, Ernest (1993). Review: Proper Functionalism and Virtue Epistemology. *Noûs*. vol. 27, no. 1, pp. 51-65.

- (1999). How to Defeat Opposition to Moore. *Noûs*. vol. 33. Supplemental issue 13: Epistemology. pp. 141-153.
- (2002). The Place of Truth in Epistemology. In Michael DePaul and Linda Zagzebski, eds. *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Clarendon Press; New York: Oxford University Press.
- (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume 1*. Oxford: Clarendon Press; New York: Oxford University Press.
- Steglich-Petersen, A. (2011). How to be a Teleologist about Epistemic Reasons. In A. Reisner and A. Steglich-Petersen, eds. *Reasons for Belief*. Cambridge: Cambridge University Press.
- Tannsjo, Torbjörn (2010). *From Reasons to Norms: On the Basic Question in Ethics*. Dordrecht: Springer.
- Taylor, Shelly, and Brown, Jonathan (1988). Illusion and Well-Being: A Social Psychological Perspective on Mental Health. *Psychological Bulletin*. vol. 103, pp. 193-210.
- (1994). Positive Illusion and Well-Being Revisited: Separating Fact from Fiction. *Psychological Bulletin*. vol. 116, pp. 21-27.
- Turri, John (2010). On the Relationship between Propositional and Doxastic Justification. *Philosophy and Phenomenological Research*. vol. 80, no. 2, pp. 312-326.
- (2011). Believing for a Reason. *Erkenntnis*. vol. 74, no. 3, pp. 383-397.

- Vahid, Hamid (2003). Truth and the Aim of Epistemic Justification. *Teorema*. vol. 22, no. 3, pp. 83-91.
- (2006). Aiming at Truth: Doxastic vs. Epistemic Goals. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. vol. 131, no. 2, pp. 303-335.
- Velleman, David (2000). *The Possibility of Practical Reason*. New York: Oxford University Press.
- Wedgwood, Ralph (2002). The Aim of Belief. *Philosophical Perspectives*. vol.16: Language and Mind, pp. 267-297.
- White, R. (2007). Epistemic Subjectivism. *Episteme: A Journal of Social Epistemology*. vol. 4, no. 1, pp. 115-129.
- Williams, Bernard. (1981) Internal and External Reasons. In his *Moral Luck*. Cambridge: Cambridge University Press, pp. 101-113.
- Wood, Allen (2008). The Duty to Believe According to the Evidence. *International Journal for the Philosophy of Religion*. vol. 63, pp. 7-24.
- Wrenn, C. B. (2004). Hypothetical and Categorical Epistemic Normativity. *The Southern Journal of Philosophy*. vol. 42, pp. 273–290.
- Zagzebski, Linda (2009). *On Epistemology*. Belmont, CA: Wadsworth.