

## THE AIM OF BELIEF<sup>1</sup>

Ralph Wedgwood  
Merton College, Oxford

### 0. Introduction

It is often claimed that beliefs aim at the 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. First, I shall explain how I understand this normative claim, and then I shall try to make it plausible that it expresses a fundamental truth about belief.<sup>2</sup> In the course of doing this, I shall also give a sketch of an account both of rational belief and of knowledge.

### 1. Normative concepts

One might wonder whether the claim that I am focusing on—roughly, the claim that a belief is correct if and only if the proposition believed is true—is *trivial*. It certainly would be trivial if ‘belief’ here just meant “proposition that is believed”, and ‘correct’ were just a synonym for ‘true’. But as I am using the term here, a “belief” is not just a proposition that is believed; it is a *particular mental state* that a person has or forms on a particular occasion. ‘Correct’ is also not just a synonym for ‘true’. To say that a mental state is “correct” is to say that in having that mental state, one has got things “right”; one’s mental state is “appropriate”. To say that a mental state is “incorrect” is to say that in having that mental state, one has got things “wrong” or made a “mistake”; one’s mental state is in a sense “defective”. Clearly, there is nothing wrong or defective about false propositions as such; what is defective is *believing* such false propositions. Moreover, other mental states besides beliefs, such as choices or

decisions, can also be wrong or mistaken or incorrect. So ‘is correct’ also does not just mean “is a belief in a true proposition”.

As I am using it here, the term ‘correct’ expresses a *normative concept*. I cannot give a full analysis here of what it is for a concept to be a normative concept. However, I will propose a *sufficient* condition for normativity.<sup>3</sup> (In this paper, I shall be concerned only with normative concepts that meet this sufficient condition for normativity.)

I propose that certain concepts are normative because it is a *constitutive* feature of these concepts that they play a *regulative* role in certain practices. Suppose that a certain concept ‘*F*’ is normative for a certain practice. Then it is a constitutive feature of the concept ‘*F*’ that if one engages in this practice, and makes judgments about which moves within the practice are *F* and which are not, one is thereby *committed* to regulating one’s moves within the practice by those judgments. Perhaps, for example, if one engages in this practice, and makes a judgment about moves that are available to one, of the form ‘Move *x* is *F* and move *y* is not *F*’, one is thereby committed to making move *x* rather than move *y*, if one makes either. For instance, the concept of a “legal chess move” seems to be normative for the ordinary practice of playing chess in this way.

What does it mean to say that engaging in this practice and making the judgment ‘Move *x* is *F* while move *y* is not’ “commits” one to not making move *y*? Roughly, it means that it is *irrational* for one simultaneously to engage in this practice, to make the judgment ‘Move *x* is *F* while move *y* is not’, and yet to make move *y*.<sup>4</sup> Making move *y*, while engaging in the practice and making this judgment, is “irrational” in the sense that it involves having an incoherent set of mental states—a set of mental states that intuitively conflict with each other. For example, engaging in the “ordinary practice of playing chess” presumably involves aiming to win a game of chess by making only legal moves. So, making what one judges to be an illegal move, while engaging in the ordinary practice of playing chess, involves a set of mental states—the aim of not making any illegal moves, the judgment that *y* is an illegal move, and the decision to make move *y* anyway—that intuitively conflict with each other.

If the concept ‘*F*’ is normative for a certain practice in this way, then engaging in the practice commits one to treating the judgment that a certain move is *not F* as representing a *decisive reason* against making that move (at least if there is an available alternative move that one judges to be *F*). To say that there is a “decisive reason” for one not to make a certain move is to say that one (in some sense) “ought not” to make that move. So, engaging in that practice commits one to accepting that one (in some sense) “ought not” to make moves within the practice that are not *F*.<sup>5</sup>

In what follows, I shall focus on a particularly fundamental “practice”—namely, in the broadest sense, *reasoning*. The “moves” within this practice in-

clude forming and abandoning beliefs (or, in the case of practical reasoning, choices). I shall assume that one cannot make a normative judgment about any of these “moves” without engaging in the “practice” of reasoning. So from now on, I shall not have to mention that one is engaging in this “practice” when one makes such judgments about these “moves”.

It seems that there are at least two concepts that are normative for the practice of reasoning—namely, the concepts ‘rational’ and ‘correct’. For example, suppose that you judge that it is rational for you to suspend judgment about  $p$  and not rational for you to believe  $p$ . Then it is a constitutive feature of the concept ‘rational’ that you are thereby committed to not believing  $p$ . In effect, if you make judgments about what is and what is not rational for you to believe, you are thereby committed to accepting that you (in some sense) “ought not” to hold beliefs that are not rational.<sup>6</sup> Similarly, it is a constitutive feature of the concept ‘correct’ that, if you judge that it is correct for you to disbelieve  $q$  and not correct for you to believe  $q$ , you are thereby committed to not believing  $q$ . If you make judgments about what it is correct for you to believe and what it is not, you are thereby committed to accepting that you (in some sense) “ought not” to hold beliefs that are not correct.<sup>7</sup>

## 2. Epistemic norms

In this way then, the concept ‘correct’ (like the concept ‘rational’) is normative for the practice of reasoning. But what distinguishes the concept ‘correct’ from other such normative concepts? To answer this question, I shall have to introduce a new notion—the notion of “epistemic norms”.

In general, all reasoning consists of *revising* one’s mental states, for some reason or other. (I shall use the term ‘revise one’s mental states’ broadly, so that it includes not just forming a new mental state but also abandoning or reaffirming an old mental state.) There are two fundamentally different kinds of reasoning. One kind is *theoretical reasoning*, which consists of revising one’s beliefs; the other kind is *practical reasoning*, which consists of revising one’s choices or intentions.

Some normative concepts are normative for the practice of practical reasoning, but not for the practice of theoretical reasoning. For example, consider the concept of “having disastrous consequences”. Suppose that one judges ‘My believing  $p$  would have disastrous consequences (while my not believing  $p$  would not have disastrous consequences)’. Intuitively, this judgment does not rationally commit one to not believing  $p$ ; at most, it commits one to intending to try to bring it about that one does not believe  $p$ . There need be nothing irrational about one’s simultaneously judging ‘It would have disastrous consequences for me to believe  $p$ ’ and yet believing  $p$ . One might make this judgment, and attempt to bring it about that one does not believe  $p$ , but fail in this attempt (perhaps because one cannot make oneself forget the overwhelmingly powerful

evidence in favour of  $p$ ). This might be unfortunate, but it need not be in any way irrational; it need not involve having an incoherent set of mental states—that is, a set of mental states that intuitively conflict with each other.<sup>8</sup>

On the other hand, consider the judgment ‘It would be incorrect for me to believe  $p$ ’. This judgment commits one, not merely to trying to bring it about that one does not believe  $p$ , but directly to not believing  $p$ . If one judges ‘It would be incorrect for me to believe  $p$ ’, but fails in one’s attempt to bring it about that one does not believe  $p$ , this would not only be unfortunate—it would be irrational. So the concept of “being incorrect” (unlike the concept of “having disastrous consequences”) really is normative for the practice of theoretical reasoning.

Take any concept ‘ $F$ ’ that is in this way normative for the practice of theoretical reasoning. Suppose that it is true that some belief  $b$  is  $F$ . It seems plausible that all such normative truths *strongly supervene* on truths that can be stated using only non-normative concepts.<sup>9</sup> That is, it is impossible for there to be a belief that is exactly like  $b$  in all non-normative respects that is not also  $F$ . So,  $b$  must have some property  $A$ —a property that can in principle be ascribed using only non-normative concepts—such that it is a necessary general principle that all beliefs that have property  $A$  are  $F$ . I shall call all such necessary general principles “epistemic norms”.

There are reasons for thinking that these epistemic norms articulate *essential* features of the types of beliefs that they apply to. First, as I have defined the term, “epistemic norms” are always necessary truths; so in that sense, these norms articulate *necessary* features of the types of beliefs that they apply to. Second, it seems plausible that, quite generally, types of mental state are *individuated* by the conditions under which they satisfy normative concepts. Suppose that there were two distinct types of mental state that did not differ in any way with respect to the conditions under which they satisfy normative concepts. Then these two types of mental state would count as correct under exactly the same conditions; they would count as rational under exactly the same conditions; they would be supported by exactly the same reasons, and would constitute reasons for exactly the same further actions and attitudes, under all possible circumstances; and so on. But in that case, it is very hard to see how these really could be distinct types of mental state at all. Moreover, every attempt to individuate types of mental state in purely non-normative terms seems to face serious problems.<sup>10</sup> For these reasons, it is plausible that epistemic norms apply to types of belief in virtue of the very *essence* or *nature* of those types of belief. If there are some epistemic norms that apply to absolutely *all* beliefs (just to have a label, we could call them “universal epistemic norms”), then these universal epistemic norms would apply to beliefs in virtue of the very essence or nature of *belief as such*.

Presumably, some of these universal epistemic norms are just *primitive* truths that cannot be any further explained, while the other universal epistemic norms (if any) are ultimately explained on the basis of some such primitive truths (perhaps including some of these “primitive” epistemic norms). If I am right to

suggest that these universal epistemic norms apply to beliefs in virtue of the very essence or nature of belief, it would seem that these “primitive” norms and truths actually articulate that essence or nature of belief. They would articulate, as we might put it, *constitutive* features of belief—that is, features that make belief the type of mental state that it is.<sup>11</sup>

### 3. Correct belief and the fundamental epistemic norm

Suppose that there is a universal epistemic norm—that is, a norm that applies to all beliefs as such—that is particularly *fundamental*, in the following way. Not only is this a “primitive” epistemic norm that cannot be any further explained; but it also explains absolutely all other such universal epistemic norms. If there is such a norm, then, I propose, a belief is “correct” just in case it satisfies this fundamental epistemic norm.

According to my proposal, this is precisely what is distinctive of the concept ‘correct’, in contrast to all other normative concepts. A belief counts as “correct” just in case it satisfies the most *fundamental* of all universal epistemic norms. (This approach could be applied to choices as well as to beliefs. According to this approach, for a choice to be “correct” would be for it to satisfy the most fundamental of universal *practical* norms—where “practical norms” are necessary principles that specify when choices satisfy the concepts that are normative for the practice of *practical* reasoning.)

What might it mean to say that all the other universal epistemic norms are “explained by” such a fundamental epistemic norm? I shall focus on two such universal epistemic norms here—the norm that specifies when beliefs count as “rational”, and the norm that specifies when beliefs count as “knowledge”. I have already argued that the concept of “rational belief” is normative for the practice of theoretical reasoning. It is a constitutive feature of this concept that if one judges that a certain belief would not be a “rational” belief for one to hold, this judgment commits one to not holding that belief. In the final section of this paper, I shall argue that something similar also holds of the concept ‘knowledge’. So the necessary general principles that specify when beliefs count as rational, and when beliefs count as knowledge, both count as “universal epistemic norms”.

As I shall understand it, to say that these universal epistemic norms of rational belief and of knowledge are “explained by” the fundamental epistemic norm of correct belief is to say the following. Suppose that this universal norm of rational belief is the principle that all and only beliefs that have property *R* are rational; and suppose that this universal norm of knowledge is the principle that all and only beliefs that have property *K* count as knowledge. Then this fundamental norm of correct belief, when conjoined with various other truths that are not themselves epistemic norms, implies that there must be an epistemic norm that requires beliefs to have property *R*, and also that there must be an epistemic norm that requires beliefs to have property *K*.

My central proposal in this paper is that there is such a fundamental epistemic norm of correct belief—to put it roughly, the principle that a belief is correct if and only if the proposition believed is true. However, this formulation still needs further refinement. A fully adequate statement of this norm must apply to *all* types of belief. As we might say, this norm must apply to all “doxastic” or “credal” attitudes that one might have towards any proposition.

For example, many philosophers believe that some of the “doxastic” or “credal” attitudes that one might have towards a proposition are “partial beliefs”—that is, states in which one puts some credence in the proposition, but also puts some credence in the proposition’s negation. If this is right, then the fundamental epistemic norm of correct belief must be reformulated to apply to such “partial beliefs” as well. The most promising way to reformulate this norm so that it applies to such partial beliefs is as follows. If a proposition is true, then the *higher* one’s degree of credence in the proposition, the *closer* the belief is to being correct; and if the proposition is not true, the *lower* one’s degree of credence, the closer the belief is to being correct.<sup>12</sup> But for the most part, I shall ignore these partial beliefs here. I shall chiefly focus on outright, all-or-nothing beliefs—states in which one simply believes a proposition, giving no thought at all to how much credence to put in the proposition’s negation. The task of extending my approach to such partial beliefs must await another occasion.

Even if we set these “partial beliefs” aside, however, the “doxastic” or “credal” attitudes that one might have towards a proposition *p* include, not only the state of believing *p*, but also the state of *disbelieving p*.<sup>13</sup> To accommodate the state of disbelieving *p* as well as the state of believing *p*, we must reformulate the fundamental norm as follows. If *p* is true, then the state of believing *p* is correct, while the state of disbelieving *p* is *incorrect*—that is, it is as far from being correct as it is possible for a belief to be. On the other hand, if *p* is *not* true, the state of believing *p* is incorrect, and the state of disbelieving *p* is correct.<sup>14</sup>

Another doxastic or credal attitude that one might have towards *p*, I shall assume, is the attitude of *suspending judgment* about *p*. Suspending judgment about *p* is quite different from simply neither believing nor disbelieving *p*. (The property of neither believing nor disbelieving *p* is not a type of mental state at all—even rocks and numbers have *that* property.) As I shall use the term, one “suspects judgment” about *p* when one consciously considers *p*, but neither believes nor disbelieves *p*. (To “consider” *p* is just to “entertain” *p*; it is for *p* to “occur” to one in occurrent thinking.) Thus, while one is consciously considering *p*, one cannot help having some broadly speaking doxastic or credal attitude towards *p*: either one believes *p*, or one disbelieves *p*, or one does neither—in which case one suspends judgment about *p*.

According to the fundamental norm of correct belief, I propose, suspending judgment about *p* is neither correct nor incorrect. If one suspends judgment about *p* then one has neither got things right nor got things wrong about *p*.

Thus, this fundamental norm ascribes to the state of suspending judgment about  $p$  an intermediate value, between the complete correctness of believing  $p$  when  $p$  is true, and the complete incorrectness of believing  $p$  when  $p$  is not true.

One might wonder whether suspension of judgment about  $p$  is always inferior, from the standpoint of the fundamental epistemic norm, to believing  $p$  when  $p$  is true. What if  $p$  is an utterly tedious or trivial truth? Then wouldn't suspending judgment about  $p$  be a better way to invest one's scarce cognitive resources? In fact, however, given that one is already consciously considering  $p$ , one can believe  $p$  with hardly any investment of cognitive resources at all: one can believe  $p$  at the moment when one consciously considers  $p$ , even if one forgets  $p$  immediately afterwards. (Of course, *investigating* whether  $p$  is true may involve an enormous investment of resources. But to say that believing  $p$  when  $p$  is true is always better than suspending judgment about  $p$  is *not* to say that carrying out costly investigations to determine whether  $p$  is true is always better than suspending judgment about  $p$ .) Admittedly, if  $p$  is an utterly tedious or trivial truth, then it may be a complete waste of time even to consider  $p$ . But the fundamental epistemic norm says nothing about whether or not one should consider  $p$ , since this norm only concerns propositions that one *actually consciously considers*.<sup>15</sup> Thus, this norm simply does not compare one's believing  $p$  with the state of affairs in which one never considers  $p$  at all. This norm has nothing to say about which of these two states of affairs is better or worse than the other.<sup>16</sup> But if one *is* consciously considering whether  $p$  is the case, then, according to this fundamental epistemic norm, if  $p$  is true, believing  $p$  is better than suspending judgment about  $p$ .

I shall take it then that this principle—roughly, that a belief is correct if and only if the proposition believed is true—gives an intuitively plausible specification of what it is for beliefs to count as correct. I have already argued that the concept 'correct' is normative for the practice of theoretical reasoning; and since this principle applies to all beliefs, and is clearly necessary and not contingent, it follows that this principle is a universal epistemic norm. In the rest of this paper, I shall try to make it plausible that this principle is the *fundamental* epistemic norm—that is, that it explains the norms of rational belief and of knowledge, in the way that I have outlined.

#### 4. The norm of rational belief

According to my central proposal, the fundamental epistemic norm implies that, for every proposition  $p$  that one consciously considers, the best outcome is to believe  $p$  when  $p$  is true, the second best outcome is to suspend judgment about  $p$ , and the worst outcome is to believe  $p$  when  $p$  is false. In this section, I shall try to make it plausible that the norm of rational belief is entirely explained by this fundamental epistemic norm.

In the face of this suggestion—that the norm of rational belief is entirely explained by this fundamental epistemic norm—two objections immediately

spring to mind. First, it may seem intrinsically implausible that *practical considerations*—such as one’s needs and interests and so on—play absolutely no role in determining whether or not one rationally should believe a proposition. Of course, the proponent of this suggestion may reply that such practical considerations may play an extensive role in determining which propositions one should *consider* in the first place. But according to this second objection, that is not enough: even given that one is considering this proposition, practical considerations must play a role in determining whether or not one rationally should believe that proposition.

Second, it may seem that this fundamental norm does not have enough content or structure to explain the norms of rational belief. In particular, this norm says nothing to answer the questions of *how much* better it is to believe *p* when *p* is true than to suspend judgment about *p*, and *how much* better it is to suspend judgment about *p* than to believe *p* when *p* is false. But different answers to these questions have dramatically different implications about when one should believe *p* and when one should suspend judgment about *p*. If suspending judgment about *p* is *much* better than believing *p* when *p* is false, but *not* much worse than believing *p* when *p* is true, then presumably the rational attitude is to suspend judgment unless the evidence for *p* is very strong. On the other hand, if suspending judgment about *p* is *much* worse than believing *p* when *p* is true, but *not* much better than believing *p* when *p* is false, then presumably the rational attitude is to take one’s chances and believe *p* even if the evidence for *p* is relatively weak.<sup>17</sup> As we might put it, this fundamental norm does not determine how one is to balance the value of having a correct belief about *p* against the disvalue of having an incorrect belief about *p*; so it cannot determine when it is rational to believe *p* and when it is rational to suspend judgment about *p*.

The point behind this second objection is entirely correct. The fundamental epistemic norm of correct belief, as I have formulated it, does not determine any unique way of balancing the value of having a correct belief about *p* against the disvalue of having an incorrect belief about *p*. If the norm of rational belief is indeed explained by this fundamental norm, the norm of rational belief must leave it indeterminate exactly when it is rational to believe *p* and when it is rational to suspend judgment about *p*.

If the norm of rational belief is indeterminate in this way, then there will be many precisifications of the concept of “rational belief”, none of which is more faithful to the concept than any other. On some of these precisifications, the concept of rationality is quite strict: on these precisifications, there are very few propositions in which it is rational to have an outright, all-or-nothing belief; for all other propositions, the only rational attitude is to suspend judgment (or to have a mere “partial” degree of belief). On other precisifications, however, the concept of rationality is much less strict: there are many more propositions in which it is rational to have an outright belief.

This is not to say that there are no limits to which precisifications of the concept of “rational belief” count as faithful to the concept. It seems plausible

that none of these admissible precisifications will allow either of these two considerations—the value of having a correct belief about  $p$ , and the disvalue of having an incorrect belief about  $p$ —entirely to swamp the other. There is no admissible precisification on which it is rational to suspend judgment about absolutely all propositions, irrespective of how much evidence for those propositions one may have; in this way, the disvalue of having an incorrect belief never entirely swamps the value of having a correct belief. Equally, there is no admissible precisification on which it is rational to believe a proposition as a wild shot in the dark—that is, to believe a proposition even though its level of credibility is no greater than that of its negation; in this way, the value of having a correct belief never entirely swamps the disvalue of having an incorrect belief.

Nonetheless, within these limits, there is still a range of precisifications of the concept of “rational belief” that are all equally faithful to the concept. If one asks whether it is rational for someone (whether it is oneself or someone else) to hold a certain belief, the *context* in which one asks this question often narrows down the range of precisifications that are relevant to answering the question. One aspect of the context that may do this consists of the *practical considerations* that are salient in that context.

For example, imagine the following two cases. In both cases, there is a certain amount of evidence in favour of a proposition  $p$ . On a more relaxed precisification of the concept of rationality, this evidence is enough to make it rational to believe  $p$ . On a stricter precisification, however, this evidence is *not* enough to make it rational to believe  $p$ ; the only rational attitude is to suspend judgment about  $p$  (or to have a mere partial degree of belief in  $p$ ). Now suppose that these two cases differ in the practical *costs* of being wrong about  $p$ . In the first case, if  $p$  is false, then the costs of one’s having an outright belief in  $p$  are *much* higher than the costs of suspending judgment (or having a mere partial belief in  $p$ ); on the other hand, if  $p$  is true, the benefits of having an outright belief in  $p$  are *not* much greater than the benefits of suspending judgment about  $p$  (or having a mere partial belief in  $p$ ). If one asks what it is rational to believe in this first case, the relevant precisification may be the stricter of the two, according to which the rational attitude is complete suspension of judgment (or, if one needs some degree of belief in order to act, a mere partial degree of belief). In the second case, by contrast, if  $p$  is false, the costs of having an outright belief in  $p$  are *not* much higher than the costs of suspending judgment (or having a mere partial degree of belief), whereas if  $p$  is true, the benefits are much higher. If one asks what it is rational for one to believe in this second case, the relevant precisification may be the more relaxed of the two, according to which it is rational for one to have an outright belief in  $p$ .

In this way, practical considerations may indeed be relevant, in certain contexts, to answering the question ‘Is it rational for  $x$  to believe  $p$ ?’. There may be other contexts, however, in which the range of precisifications that are relevant to answering this question is determined, not by any such practical considerations, but purely by the *habits* of the thinker (or the participants in the

conversation), which make that range of precisifications salient in the context. So the question ‘Is it rational for  $x$  to believe  $p$ ?’ is not *essentially* connected to such practical considerations at all.

However, it may still seem implausible that this fundamental epistemic norm—roughly, the principle that a belief is correct if and only if the proposition believed is true—can explain the norms of rational belief. After all, according to this principle, *any* belief in a true proposition is correct—even if the belief in question is grossly irrational. So how can this principle explain the norms of rational belief?

To revert to the metaphor of “aiming at truth”, the answer is this. Even though irrational beliefs can be correct, the only way in which it makes sense to *aim* at having a correct belief is *by means of* having a rational belief.

Suppose that you are considering a proposition  $p$ , and are literally “aiming” to conform to this fundamental epistemic norm. That is, you are aiming to believe  $p$  if and only if  $p$  is true. Clearly, you will not end up achieving this aim simply because it is your aim. You will have to *do something* in order to achieve this aim. That is, you will have to do something *by means of which* (if all goes well) you will achieve this aim. Presumably, in order to achieve this aim, you must revise your beliefs in certain ways when you are in certain conditions, and revise your beliefs in other ways when in other conditions. We may imagine a set of *rules*, such that each of these rules permits one to revise one’s beliefs in a certain way whenever one is in a certain related condition. For example, one such rule might permit one to believe a proposition  $p$  whenever one has an experience or apparent perception as of  $p$ ’s being the case, and has no special reason to think that one’s experiences are unreliable in the circumstances.<sup>18</sup>

However, if it is purely a *fluke* that one conforms to these rules, it will hardly be appropriate to say, even metaphorically, that these rules are the “means” that one uses in order to pursue an aim. This description will be appropriate only if one also *follows*, or is *guided by*, these rules. We need not worry here exactly what it is to “follow a rule”, so long as it is clear that one can “follow a rule” even in forming beliefs in a completely spontaneous, unreflective way.<sup>19</sup> So, in order to achieve the goal of believing the proposition  $p$  if and only if  $p$  is true, you must revise your beliefs by means of following certain rules. But which rules does it make sense for you to follow, in order to achieve this goal?

Suppose that you are actually trying to *choose* which rules to follow, in order to achieve this goal. What is the rational way for you to choose which rules to follow? Since the furthest from this goal that you can be is to end up believing something false, you should presumably aim not to follow any rules that, in the circumstances, might easily result in your believing something false. It is not enough here for you to aim to follow rules that are merely *generally reliable*—that is, rules that yield a high ratio of true to false beliefs. If you know that certain rules, though generally reliable, could easily lead to your believing a false proposition in the circumstances at hand, then you should not follow those rules. This is not to say that you should aim not to follow any

rules that might easily lead to your believing something false in *any* circumstances. All that matters is that they could not easily lead to your believing a false proposition in the circumstances at hand—as I shall put it, these rules must be *reliably error-avoiding in the circumstances*.<sup>20</sup> Equally, you need not aim to follow rules that could not easily lead *anyone* to believe a false proposition in the circumstances; all that is necessary is that these rules could not easily lead *you* to believe a false proposition in the circumstances.

In general, we may say that a set of rules is reliably error-avoiding for the thinker in the circumstances just in case, in all nearby possible worlds, these rules do not lead to the thinker's believing a false proposition in any case in which the thinker follows the rules in those circumstances.<sup>21</sup> Something is the case in one of these “nearby possible worlds” just in case it is actually, or could easily have been, the case. Since these “nearby possible worlds” include the actual world, these rules cannot be reliably error-avoiding in the circumstances if they actually yield belief in a proposition that is false.

There seems to be considerable indeterminacy in this notion of the “nearby possible worlds”, or of what “could easily” have been the case. Perhaps the notion of the “nearby possible worlds” guarantees that these nearby possible worlds must include more than just the *actual* world, and cannot include absolutely *all* possible worlds. But between these two extremes, there are many sets of possible worlds that have equal claim to be regarded as being the “nearby possible worlds”. If a very large number of worlds count as “nearby possible worlds”, then the corresponding standard of reliability will be extremely demanding. If only a few worlds count as “nearby possible worlds”, then the corresponding standard of reliability will be much lower.

So the notion of “reliability” that figures in the thesis that I have just advanced—that in order to achieve your goal of believing *p* if and only if *p* is true, it is rational for you to aim to follow only reliably error-avoiding rules—is indeterminate. Given this thesis, it follows that the question ‘Is it rational for you to aim to follow rules that meet standard of reliability *S*, in order to achieve your goal of believing *p* if and only if *p* is true?’ does not always have a determinate answer. In the same way as I suggested above for the simpler question, ‘Is it rational for you to believe *p*?’ this indeterminacy may be reduced to some degree by the *context* in which this question is considered. In some contexts, it will count as “rational” for you to aim to follow rules that meet a relatively low standard of reliability; in other contexts, it will not count as “rational” for you to aim to follow any set of rules that does not meet a relatively high standard of reliability. In general, this will be because a certain standard of reliability (or range of standards of reliability) is *salient* in those contexts—often (although not always) because of the *practical considerations* (such as the needs or purposes or values) that are salient in those contexts.

According to my definition, a set of rules is “reliably error-avoiding in the circumstances” just in case it could not easily happen, in the circumstances, that these rules would lead you to believe something false. A set of rules can be reliably error-avoiding even if it is unnecessarily restrictive: for example, a re-

liably error-avoiding set of rules might not permit one to form any beliefs at all on Tuesdays and Thursdays. But we are supposing that your aim is not just to avoid believing  $p$  if  $p$  is *not* true, but also to believe  $p$  if  $p$  is true. So you should aim not to follow rules that are unnecessarily restrictive in this way. Instead, you should aim to follow rules that are no more restrictive than is necessary in order to make it rational to regard those rules as reliably error-avoiding. That is, you should aim to follow a set of rules such that there is no second set of rules available to you that it is equally rational for you to regard as reliably error-avoiding, which differs from the first set only in that it permits one to hold the same beliefs in a wider range of conditions.

Since you are aiming, not only to avoid believing  $p$  if  $p$  is not true, but also to believe  $p$  if  $p$  is true, it is not enough merely that the rules that you are following are both reliably error-avoiding and no more restrictive than necessary. You should also aim to use these rules, if you can, in a way that will reliably yield belief in  $p$  if  $p$  is true. Let us say that your “method” for reaching belief in  $p$  consists of a partially ordered series of “steps”, where each of these steps consists in your following some rule. You should aim, if you can, to use a method such that it could not easily happen, in the circumstances, that using this method would fail to yield a belief in  $p$  when  $p$  is true. Let us say that such a method is *reliably belief-yielding in the circumstances* with respect to  $p$ . More precisely, a method is reliably belief-yielding in the circumstances with respect to  $p$  just in case, in all nearby possible worlds, the method yields belief in  $p$  in every case in which the thinker uses the method in those circumstances and  $p$  is true.

So, it seems, the rational way for you to choose which rules to follow in the circumstances is this. First, you should restrict your choice to sets of rules that it is rational for you to believe to be reliably error-avoiding in the circumstances, but also no more restrictive than necessary in order for it to be rational to regard those rules as reliably error-avoiding. Second, if you can, you should choose a set of rules that it is rational for you to believe to provide a method that is reliably belief-yielding, in the circumstances, with respect to the proposition that is in question. For short, let us say that you should choose rules that it is rational for you to regard as “sufficiently reliable in the circumstances”.<sup>22</sup>

This is *not* to say that the rational way to choose which rules to follow in the circumstances is to choose rules that are *in fact* sufficiently reliable in the circumstances. Even if you are bedevilled by an evil demon who ensures that whatever rules you choose, those rules will lead to your believing a false proposition in the circumstances, it may still be *rational* for you to regard certain rules as sufficiently reliable in the circumstances. If so, then it would also be rational for you to choose to follow those rules in the circumstances. Conversely, a rule might actually be reliable in the circumstances, even though it is not rational for you to regard it as reliable. To take an example from Lawrence BonJour (1980), suppose that you have a perfectly reliable clairvoyant power, which always gives you reliably correct beliefs about far-away events. In this case, the rule that permits you to form the beliefs that are suggested by this

clairvoyant power is a perfectly reliable rule. However, it may be that it is not rational for you to regard this rule as reliable; and in that case, it would not be rational for you to choose to follow this rule.

Now perhaps we can claim that, in any circumstances, the rules that it is actually rational for one to follow are exactly the same as the rules that it is rational for one to *choose* to follow in those circumstances. Then, given what I have just argued, we could conclude that the rules that it is rational for one to follow in the circumstances are all and only those rules that it is rational for one to believe to be sufficiently reliable in the circumstances.<sup>23</sup>

However, this conclusion is open to a serious objection. Offhand, the conditions that make it rational to follow a set of rules seem *weaker* than the conditions that make it rational to believe those rules to be reliable.<sup>24</sup> Surely, rationally believing that a rule is reliable is a much more sophisticated achievement than simply rationally following the rule?

As it stands, this objection is unfair. This conclusion does not imply that, for it to be rational for a thinker to follow a rule, the thinker must actually *believe* that the rule is reliable: all that is necessary is that it be *rationaly permissible* for the thinker to hold such a belief. Clearly, it can be rationally permissible for a thinker to believe something, even if she does not actually believe it. But the underlying point behind the objection still needs an answer. For it to be rational for one to believe that a rule is reliable, it seems that one must have the ability to pick out or think about the rule. But isn't that too a highly sophisticated achievement?

In fact, the conclusion under consideration only requires that it must be rational for one to have a *de re* belief about the rule—that is, to believe, of the rule in question, that it is reliable (compare Audi 1993, Essay 8). To have a belief of this kind, one only needs some way of picking out what is in fact the rule. One need not be able to give a precise analysis of the essential structure of the rule.<sup>25</sup> For example, one could pick out the rule that one is actually following or is inclined to follow demonstratively, as “thinking in *this* sort of way”. Still, it is hard to see how it could be rational for one to believe a rule to be reliable unless one is in a position to grasp at least roughly what sort of rule it is. Perhaps adult human thinkers are normally in a position to acquire some such grasp of the rule that they are following, simply by reflection. But couldn't there be exceptions to this—cases in which a thinker rationally follows a rule but is not in a position to grasp what sort of rule it is?<sup>26</sup>

However, if we restrict ourselves to cases in which the thinker *is* in a position to grasp at least roughly what sort of rule he is following, then the conclusion under consideration seems plausible. If one is in a case of this sort, then it is rational to follow the rule just in case it is rational for one to believe the rule to be sufficiently reliable in the circumstances.

Moreover, it is plausible that none of these rules applies *only* to cases in which one is not in a position to grasp what sort of rule one is following. In general, what makes it rational for a thinker to follow one of these rules is a certain *relation* that holds between the thinker and the rule. (For example, per-

haps it is rational for you to follow the rule because you possess some concept such that it is constitutive of possessing this concept that one has some mastery of this rule, or because in some other way the rule plays a “pivotal” role in your cognitive economy.)

This suggests that we could define the notion of its being “rational for one to believe the rule to be reliable” in the following way. According to this definition, to say that it is “rational for one to believe rule *R* to be reliable” is to say that one stands in a certain relation to the rule *R*, such that it is rational for any thinker who grasps what sort of rule *R* is to respond to the fact that he stands in that relation to *R* by believing *R* to be reliable in the circumstances. Understood in this way, it could be rational for one to believe a rule to be reliable even if one is not in a position to grasp what sort of rule it is—indeed, even if one is incapable of having beliefs about the rule at all. So I propose that we should accept the conclusion that I suggested—that it is rational for one to follow a rule just in case it is rational for one to believe the rule to be reliable—when it is understood in this way.

Some philosophers may object that this conclusion will generate a vicious regress. According to this conclusion, it is not rational to follow a rule unless it is rational to believe the rule to be reliable. But what can make it rational to believe that rule to be reliable, if not a further rule that it is also rational to follow—and so too, according to this conclusion, rational to believe to be reliable? However, there is no reason to agree with this objection that the only thing that can make it rational to believe a rule to be reliable is a *further* rule. It may be that certain rules are just *basic*, in the sense that it is rational to follow these rules, even though the only way in which one can reach a rational belief in those rules’ reliability is by means of following those very rules.<sup>27</sup> It is still a *constraint* on the rationality of following these basic rules that it must be rational to regard them as reliable. But it is not true of these basic rules (as perhaps it is of other, non-basic rules) that the only thing that *makes* it rational to follow these basic rules is the fact that it is rational to regard them as reliable. On the contrary, the rationality of following these basic rules is part of what makes it rational to regard these rules as reliable.

At all events, I propose that this conclusion gives a correct account of what it is for one to revise one’s beliefs in a rational way. One revises one’s beliefs in a rational way just in case one revises one’s beliefs through following rules that it is rational for one to follow—that is, rules that it is rational for one to believe to be sufficiently reliable in the circumstances.

This proposal can explain several important features of the rules that it is rational for one to follow in revising one’s beliefs.<sup>28</sup> For example, typically, if it is *generally rational* to follow a certain rule—that is, if it is rational to follow the rule in *every* case to which the rule applies—then the rule must contain a clause requiring the *absence of defeating conditions*. For example, many of the rules that permit one to form beliefs on the basis of one’s experience require that one should not be in conditions that make it irrational for one to regard one’s current experience as reliable. Following such a rule typically in-

volves coming to believe a proposition  $p$  in response to one's having an experience as of  $p$ 's being the case. But if one is in certain special conditions, which make it irrational for one to believe that one's experience is reliable, then following the rule will *not* involve coming to believe  $p$ . If the rule did not require the absence of defeating conditions in this way, then whenever one is in those defeating conditions, it would not be rational to regard the rule as reliable in the circumstances; so it would not be generally rational to follow the rule. As it is, however, the rule does require the absence of such defeating conditions. So, it may be that in every case to which the rule applies, it is rational for one to regard this rule as reliable in the circumstances, even if one is in the relevant defeating conditions. Thus, it may be generally rational to follow the rule.<sup>29</sup>

So far in this section, I have only dealt with rational *belief revisions*. But after one has initially formed a belief, one may hold the belief for a long time without in any way revising (or reconsidering or reaffirming) that belief. In such cases, one often forgets exactly how one came originally to form the belief. The belief just becomes part of one's stock of background beliefs. Even if it is true that the rational way to *revise* one's beliefs is by following the appropriate rules, it does not seem plausible that a rational thinker holds her *background beliefs* by means of following rules. So how can we extend the approach that has been developed here so that it applies to background beliefs as well as to belief revisions?

The solution to this problem, I believe, is to realize that background beliefs themselves play a role that is very similar to rules. We rely on these background beliefs in making belief revisions in something like the same way that we rely on rules. In general, background beliefs play at least two different roles in structuring how we revise our beliefs. First, one may simply reconsider and reaffirm a background belief—which itself counts as a “belief revision”, in the broad sense in which I am using that term. Second, background beliefs may play a crucial role in revisions of certain *other* beliefs. For example, a background belief may be used as an “inference ticket” that takes one from certain premisses to a certain conclusion, or it may form part of the information that one takes for granted in evaluating certain other propositions. For each background belief, then, there are at least two corresponding rules, one for each of these two different ways in which one may rely on the background belief in making belief revisions.<sup>30</sup>

Since background beliefs are akin to rules in this way, when we ask whether a background belief is rational, we are in effect considering the belief as amounting to the thinker's being disposed to follow a certain rule. In some contexts, we are in effect considering the belief as amounting to the thinker's being disposed to follow the *first* rule—the rule that permits the thinker to reaffirm the background belief in question. In other contexts, we are in effect considering the belief as amounting to a disposition to follow the *second* rule—the rule that permits the thinker to rely on that background belief in revising her beliefs in other propositions in the relevant way. Given the assumption that the background belief amounts to the thinker's being disposed to follow a certain rule,

we may say that the background belief is rational just in case it is rational for the thinker to follow that rule.

As I have proposed, it is rational for one to follow a rule just in case it is (in the sense that I have outlined) rational for one to believe the rule to be reliable. For example, consider the *first* rule—the rule that permits one to reaffirm one’s background belief if the relevant question arises. According to my account of reliability, this rule will be reliable in the circumstances just in case in every nearby possible world in which one holds that very token background belief, if one considers reaffirming that belief in the circumstances, one reaffirms the belief, and the proposition that one thereby believes is true. But what is it for a token background belief in some non-actual possible world to count as the “same token belief” as one that exists in the actual world? Roughly, I propose the following. A token background belief in another possible world counts as the same token belief as one that exists in the actual world just in case the belief is not only held by the same thinker and has the same (or similar) content, but also has a sufficiently similar history in the thinker’s mental life (the belief is formed and maintained in a sufficiently similar way, and so on). So, roughly, given the assumption that a background belief amounts to a disposition to follow a rule of this first sort, the belief is rational just in case it is rational for one to believe that the belief was formed and maintained in such a way that it could not easily happen in the circumstances that a belief formed and maintained in that way would be false.

Taken together, the proposals that I have made in this section give a specification of the property *R* such that the universal norm of rational belief is the principle that all and only beliefs that have property *R* are rational. Roughly, rational beliefs are beliefs that either result from, or (in the case of background beliefs) amount to, one’s following a rule or set of rules that it is rational for one to believe to be reliable. Admittedly, this specification of this property is not fully non-circular, since it uses the term ‘rational’. But for our purposes we do not need a fully non-circular specification. This specification is enough to show that the universal norm of rational belief is *explained* by the fundamental norm of correct belief.

The following principle seems a plausible claim about norms in general (not just epistemic norms). If there is a fundamental norm that directs one to achieve a certain outcome, and that outcome is an end that one can achieve only by using *means* to that end, then there is also a secondary norm that directs one to use means that it is rational for one to believe to be sufficiently reliable means to that end.<sup>31</sup> As we have seen, for every proposition *p* that one is consciously considering, the fundamental epistemic norm of correct belief directs one to believe *p* if and only if *p* is true. But this outcome is an end that one can achieve only by using means to that end; and the only available means to that end is to revise one’s beliefs by following appropriate rules. So there must also be a secondary norm that directs one to revise one’s beliefs by following rules that it is rational for one to believe to be sufficiently reliable means to that end. According to my proposal, this secondary norm is precisely the

universal norm of rational belief. In this way, the norm of rational belief is explained by the fundamental norm of correct belief.

### 5. The conditions for knowledge

In this section, I shall propose an account of knowledge. To revert to the metaphor of the “aim of belief”, one knows something when one has, through one’s own rational efforts, succeeded in achieving the aim of believing the truth and nothing but the truth. Not every case of succeeding counts as “succeeding through one’s efforts”. This distinction can be illustrated by an example that is discussed by Donald Davidson (1980, 78). Suppose that my aim is to kill you. I try to kill you by shooting you with a gun, but I miss. However, my shot causes a herd of wild pigs to stampede and trample you to death. In one sense, I have succeeded: I have brought it about that you are now dead. But I have not succeeded through my own efforts: you are not dead because everything went according to my plan.

In general, one “succeeds through one’s own efforts” just in case one succeeds precisely because the means that one used in order to succeed worked just as they were supposed to. As I proposed in the previous section, a rational thinker is using “means” in order to achieve the aim of believing  $p$  (the proposition in question) if and only if  $p$  is true. These means are the rules that she was following in revising her beliefs in  $p$ . These means “worked as they were supposed to” just in case everything that it must have been rational for her to believe, in order for it to be rational for her to make that belief revision through following those rules, really was the case. For short, let us say that the rules “worked as they were supposed to” just in case all of the “rational presuppositions” of the thinker’s making that belief revision through following those rules were true.

As I have argued, the main “rational presupposition” of one’s making a belief revision through following certain rules is that these rules are sufficiently reliable in the circumstances. So these rules “worked as they were supposed to” only if these rules really were sufficiently reliable in the circumstances. But there may also be certain other “rational presuppositions” of one’s making a certain belief revision through following certain rules. For example, consider a *non-basic* rule  $R$ , which it is rational for one to follow only because it is rational for one to believe, as a result of following *other* rules, that rule  $R$  is reliable. For instance,  $R$  might be a rule that it is rational for one to follow only because it is rational for one to trust a certain expert who has told one that rule  $R$  is reliable. For it to be rational for one to revise one’s beliefs through following rule  $R$ , it must be rational to believe, not only that  $R$  is reliable, but also that the rule of trusting this expert is reliable. In this case, rule  $R$  “worked as it was supposed to” only if both  $R$  and the rule of trusting this expert were indeed reliable.

So I propose the following definition: knowledge is a rational belief that results from the thinker’s following (or, in the case of background beliefs,

amounts to the thinker's being disposed to follow) a rule or set of rules that "worked as it was supposed to". We do not need to add that the proposition believed is true. If the rules worked as they were supposed to, they must have been reliable in the circumstances.<sup>32</sup> Since the "nearby possible worlds" always include the actual world, a rule cannot have been reliable in the circumstances if it actually yielded belief in a proposition that was false.<sup>33</sup>

Is this intuitively correct as an account of knowledge? Let us consider some of the famous cases that were inspired by Edmund Gettier (1963). If I am in an area that contains numerous barn façades, the rule that leads me to believe 'That is a barn', on the basis of an apparent perception of a barn, is not reliably error-avoiding in the circumstances. In one of the nearby worlds in which I follow precisely that rule in those circumstances, I believe something false, since my perception is in fact of a mere barn façade, not a real barn. Since this rule is not reliable in the circumstances, the belief does not count as knowledge.<sup>34</sup> Again, suppose that I rationally come to believe a false proposition,  $p$ , and rationally infer from  $p$  the true disjunctive proposition ' $p$  or  $q$ '. My belief in the disjunctive proposition is correct and rational, but it is not knowledge. This is because I hold this belief as a result of following at least two rules—the rule that leads to the belief in  $p$ , and the rule by which I infer from  $p$  to the disjunctive proposition ' $p$  or  $q$ '; and the first of these rules is not reliably error-avoiding in the circumstances.

In general, if one's belief in the conclusion of an argument is based on one's coming to believe the lemmas of the argument, then one believes the conclusion as a result of following a set of rules that also led to one's beliefs in the lemmas. One's belief in the conclusion is rational only if one's belief in each of the lemmas is also rational. Moreover, if this whole set of rules "worked as they were supposed to", then one must have come to believe each of the lemmas through following rules that "worked as they were supposed to". This is why if one's belief in the conclusion of the argument counts as knowledge, one's beliefs in all the lemmas of the argument must also count as knowledge. It must be emphasized, however, that this point only applies to *lemmas* strictly speaking—that is, *intermediate conclusions* that one came to believe *en route* to coming to believe the ultimate conclusion of the argument. This point does not apply to all the *background beliefs* on which one relied in coming to believe that conclusion. As I suggested in the previous section, many of these background beliefs play a role that is more akin to rules than to lemmas. So long as it is rational and reliable to rely on these background beliefs in the relevant way, it is not even necessary that all these background beliefs should be *true*. One can come to know something by relying on a large mass of background beliefs, even if a few of those background beliefs are false. One can also come to know something (such as the date of an eclipse) by inferring it from a scientific theory (such as Newtonian mechanics) that is strictly speaking false—so long as the theory is close enough to the truth that it is sufficiently reliable to rely on the theory in this way.<sup>35</sup> Knowledge is incompatible

with false lemmas; but it is compatible with at least certain sorts of reliance on false background beliefs.

In the previous section, I argued that the “rational presuppositions” of one’s forming a belief through following certain rules include, not just that the rules in question are reliably *error-avoiding*, but also that the method that one is using is reliably *belief-yielding*. Thus, given my account of knowledge, when one comes to know something, one must not only have followed rules that really were reliably error-avoiding; one must also have used a method that really was reliably belief-yielding. Let us consider a simple example, where one’s method consists of following a single rule just once. As I explained, many rules that permit one to form or reaffirm a belief require the absence of certain *defeating conditions*. Suppose that it is a pure fluke that you are not in such defeating conditions. Suppose, for example, that your environment is rife with defeating evidence, but by a fluke you never become aware of this defeating evidence. To take an example from Gilbert Harman (1973, 143–44), imagine that you are visiting a country that is ruled by a dictator. You come to believe a proposition *p* that you read in an early edition of a newspaper, and never become aware of the strong defeating evidence that the dictator succeeds in planting in later editions of the newspaper, and in other news reports all over the world. (It makes no difference to this case whether the defeating evidence requires believing the negation of *p*, or merely suspension of judgment regarding *p*. Either way, you would be in the defeating conditions that are specified in the rule.) In these circumstances, your method is not reliably belief-yielding: in some of the nearby possible worlds in which the proposition is true, and you use this method in the circumstances, you are in the relevant defeating conditions, and so do not form the belief in question. Thus, your belief in this case does not count as knowledge.<sup>36</sup>

At the same time, this approach does not have the false implication that one cannot know anything through a sheer fluke. To take an example from Robert Nozick (1981, 193), suppose that the bandit Jesse James is riding past. By the sheerest fluke, you are looking in the right direction when, by another fluke, his mask slips and you recognize him. In this case, it is a fluke that you know that it was Jesse James who was riding by. But this is because, in practically all the relevant nearby worlds, the opportunity for following this rule does not even arise. The rule permits one to form a belief of the form ‘That is Jesse James’ in response to an experience of the appropriate kind; and in most of the nearby worlds, you have no experience of that kind. But in each of the nearby possible worlds in which the proposition is true, and you do have an experience of the relevant kind and follow the rule in those circumstances, you form the belief that Jesse James is riding by. Thus, your method is belief-yielding in the circumstances. This case is different from the preceding case because there is a fundamental difference between never having an opportunity to follow a rule, and following the rule but (because one is in the relevant defeating conditions) refusing to form the belief in question.

According to my proposal, it must be rational to believe a rule to be reliable, if it is to be rational for one to follow the rule, and the rule must actually be reliable, if any belief that results from one's following the rule is to count as knowledge. The sort of "reliability" that I invoked in this proposal was *local* reliability, or *reliability in the circumstances*, not *global* or *general* reliability. However, it will typically not be rational for one to regard a rule as locally reliable in the particular circumstances at hand, unless it is *also* rational for one to regard the rule as reliable in a wider range of circumstances as well. Indeed, it may not be rational for one to regard the rule as reliable in the circumstances unless it is also rational for one to regard the rule as belonging to a wider class of reliable rules. In such cases, the "rational presuppositions" of one's forming a belief through following the rule include, not just that the particular rule that one followed is reliable in the particular circumstances at hand, but also that the wider class of rules is reliable in the wider range of circumstances. According to my account of knowledge, knowledge requires the truth of all these "rational presuppositions".

This point can be illustrated by an example that is due to Christopher Peacocke (1986, 145). Suppose that you rationally believe that your body is cold, on the basis of a sensation of feeling cold. Suppose further that it is true that your body is cold, but your feeling cold is not caused by your body's being cold. Your sensations are caused by neuro-scientists, who are keeping your body in a vat. However, for some reason it could not easily happen that these neuro-scientists would give you the sensation of feeling cold unless your body really was cold. Suppose that it also could not easily happen that you would have this sensation, but not form the belief 'My body is cold'. In this case, the particular rule that permits you to judge 'My body is cold' on the basis of this sensation is reliable in the circumstances. But it could not be rational for you to regard this rule as reliable unless it was also rational for you to regard a wider class of rules that permit one to form beliefs about the state of one's body on the basis of such sensations as reliable. In this case, this wider class of rules is not reliable, since your sensations are produced by the neuro-scientists, not by the state of your body. For this reason, your belief that your body is cold does not count as knowledge.

Much more would have to be said in order to show that this approach does indeed give an adequate account of knowledge. Since my approach appeals to the notion of "reliable rules" it will at some point have to deal with the "generality problem" (see Conee and Feldman 1998). Some rules are more general than others. For example, the rule "For any proposition  $p$ , believe  $p$  whenever you have an experience as of  $p$ 's being the case" is more general than the rule "Believe that it is raining whenever you have a visual experience as of rain splashing on the ground outside the window". In some circumstances, one of these rules might be more reliable than the other. If you believe, on the basis of experience, that it is raining, then to determine whether this belief counts as knowledge, we will need to know which of these two rules you are following on that occasion. But which rule is that?

The question of how to determine which rules a thinker is following in forming a given belief is undoubtedly a profound and difficult problem. However, it is not just a problem for theories that appeal to “reliable rules”. On any plausible theory, we need to know what rules the thinker is following in order to tell whether her belief is rational. It is clear, for example, that one cannot tell whether a thinker’s belief in  $p$  is rational just by knowing that (i) the belief is based on the thinker’s beliefs in  $q$  and in  $r$ , and (ii) the latter beliefs are rational. The transition from the beliefs in  $q$  and in  $r$  to the belief in  $p$  is an instance of countless many different rules. Even if one of these rules is a rule that it is rational for the thinker to follow, many more of these rules are rules that it would be utterly crazy to follow. To assess the rationality of the thinker’s belief in  $p$ , we need to know which, out of all these countless many rules, the thinker was actually following. So the “generality problem” is a problem that every theory of rational belief must face.

A second problem that any appeal to “reliable rules” must face is to say in what *range* of possible worlds and circumstances a rule must yield correct belief in order to be reliable. I have already suggested that the epistemic norms by themselves do not determine precisely what range of possible worlds and circumstances is relevant here. There are many different precisifications of the concepts of reliability, rationality, and knowledge, which are all equally faithful to the concepts themselves. Nonetheless, the precisifications of these concepts that are relevant to answering a question (either in thought or in speech) about whether or not a particular belief counts as “rational” or as a case of “knowledge” may be limited by the *context* in which this question is considered. For answering such questions, the relevant range of possible worlds and circumstances is at least roughly determined by the context—often (although not always) by the practical considerations that are most salient in that context.

This appeal to context is necessary anyway to solve a problem that would otherwise undermine my approach. Suppose, for example, that you know that I have a single ticket in a huge lottery. The chances of my winning the lottery are non-zero, but astronomically low. Given how low the chances of my winning the lottery are, it is surely rational for you to believe, simply on this basis, that I won’t win the lottery.<sup>37</sup> In forming this belief on this basis, you are following a certain rule. So, according to my account of rationality, if this belief is rational, it must also be rational for you to believe this rule to be reliable. If it is rational for you to believe this rule to be reliable, even though you know that there is a non-zero probability of my winning the lottery, the proposition that the rule is reliable must surely be consistent with the existence of such a probability. So suppose that it is true that this rule is reliable. We may presumably also suppose more generally that your circumstances are just as it must have been rational for you to assume them to be, in forming this belief in this way, so that all other “rational presuppositions” of your forming this belief in this way are true. But then, according to my account, you *know* that I won’t win the lottery. But many philosophers find it counterintuitive to say that you

“know” that I won’t win the lottery, purely on the basis of the fact that the chances of my winning are so low.

As contextualists such as David Lewis (1996, 565) have argued, however, there *are* contexts in which it is perfectly true to say that you know that I won’t win the lottery.<sup>38</sup> In these contexts, any possible worlds in which I win the lottery are just too remote to count as “nearby” possible worlds. In other contexts, however, these possible worlds count as “nearby” possible worlds. In those contexts, it is not true to say that you know that I won’t win the lottery. Moreover, since you are perfectly well aware of the chances of my winning the lottery, it is also not true in those contexts to say that it is rational for you to regard this rule as reliable. So in those contexts, it is also not true to say that it is rational for you to believe that I won’t win the lottery. In those contexts, it is only true to say that it is rational for you to have a high partial degree of belief that I won’t win the lottery. So, in fact, the degree of reliability that a rule must have, if a belief formed by following that rule is to count as knowledge, is no higher than the degree of reliability that it must be rational for the believer to believe the rule to have, if that belief is to be rational. The required degree of reliability varies with context, but it is the same in both cases.

Someone might object to this proposal, in the following way. In at least some contexts, a perfectly competent speaker might say such things as ‘The way in which I came to believe *p* may not be reliable enough for me to count as *knowing p*; but it is surely good enough to make my belief in *p* rational.’ So why doesn’t this show that the degree of reliability that it must be rational for the believer to believe the rule to have, if the belief is to be rational, is *lower* than the degree of reliability that the rule must actually have, if the belief is to count as knowledge?

In fact, however, it seems to me, there are also contexts in which a perfectly competent speaker might say ‘I know *p*, although perhaps I don’t have quite enough evidence for my belief in *p* to count as *rational*.’ (For example, suppose that the speaker is intuitively convinced of *p*, and believes that this intuitive conviction is reliable, but cannot give much in the way of a public justification of this conviction.) This seems to show that there are certain special contexts in which two different standards of reliability are salient, and by emphasizing a distinction between two epistemic terms, speakers can associate one of these terms with the stricter standard and the other term with the weaker standard. But it seems not to be built into the very meaning of these terms that one of them is associated with a stricter standard of reliability than the other.

Although context may reduce the indeterminacy in the notion of “reliability” or of the “nearby possible worlds”, it seems unlikely to eliminate it altogether. This is not in my view an objection to the account that I have proposed, since there is a corresponding indeterminacy in our intuitions about when a belief counts as knowledge. For example, suppose that you form correct beliefs on the basis of veridical experience in the ordinary way; but the evil demon had planned to fill your mind with misleading experiences today, and it

was only by the most extraordinary fluke that his plans miscarried. In this case, it is unclear whether it “could easily have happened” that following the relevant rules in those circumstances would lead to your believing something false. But it is equally unclear—or so it seems to me—whether in this case your beliefs count as knowledge. Thus, the indeterminacy in my account of knowledge does not seem to me an objection to that account.

## 6. The significance of knowledge

In this final section, I shall try to show that, if my account of knowledge is correct, then the norm of knowledge is indeed explained by the fundamental norm of correct belief—that is, roughly, by the principle that a belief is correct if and only if the proposition believed is true. To put the point crudely—reverting once again to the metaphor of “aiming”—I need to show how the fact that a person should aim to get to the truth is enough to explain why that person should aim to *know* the truth.

In fact, there are reasons for thinking that the only acceptable explanation of why we should aim at knowledge must be grounded in the more fundamental principle that we should aim at getting to the truth. A number of philosophers have recently tried to highlight the further advantages that knowledge has over mere true belief.<sup>39</sup> Such further advantages could certainly explain why we *sometimes* aim to have knowledge in addition to true belief. But what is really puzzling is that we almost *never* aim to have true belief without at the same time aiming to know. Indeed, as Bernard Williams (1978, 37) pointed out, we would not ordinarily suppose that ‘aiming to get to the truth’ and ‘aiming to know’ describe different aims at all. On the contrary, we would ordinarily assume that they describe *the very same aim*.

Broadly speaking, I believe that Williams (1978, 38–45) gave the correct account of the relation between “aiming at the truth” and “aiming to know”. The reason why we would ordinarily take these two phrases as describing the very same aim is this: a rational thinker cannot pursue the aim of believing the truth and nothing but the truth, except in a way that, if it succeeds, will result in knowledge. My account of knowledge supports this account of the connection between “aiming at the truth” and “aiming to know”. As I have already argued, a rational thinker cannot pursue the aim of believing the truth and nothing but the truth, without using means that it is rational for her to regard as reliable means to that aim. But if these means result in her believing the truth precisely because they “worked just as they were supposed to”, then (according to my account of knowledge) the belief produced by these means counts as knowledge. So there is no way for a rational thinker to pursue the truth except in a way that, if it succeeds, will result in knowledge. If this is right, then we should disagree with those philosophers, such as Christopher Peacocke (1999, 34) and Timothy Williamson (2000, 208), who suggest that belief has knowledge as one of its ultimate aims. Knowledge is not an *ultimate* aim of belief.

Belief's only ultimate aim is truth. Belief aims at knowledge only in the sense that every rational believer aims at the truth by using means that, if successful, will result in knowledge.

However, according to the approach that I am advocating here, the metaphor of "aiming at truth" should be interpreted in fundamentally normative terms. My central proposal is that the fundamental norm of correct belief explains the norm of knowledge. My account of knowledge gives a specification of that property *K* such that all and only beliefs that have property *K* count as knowledge. To show that the fundamental norm of correct belief "explains" the norm of knowledge, I must show that this fundamental norm (together with other truths that do not themselves count as epistemic norms) entails that there is a norm that requires beliefs to have property *K*. To say that there is a norm that requires that beliefs have property *K* is to say that it is necessary that whenever a belief lacks property *K*, there is a true judgment about that belief that rationally commits one to not holding that belief. So I shall try to show that the fundamental norm of correct belief explains the norm of knowledge by arguing that it is necessary that whenever one is not in a position to know *p*, then there is a judgment that one could truly make about *p*, which would rationally commit one to not believing *p*.

Suppose that you are not in a position to know *p*. Then, given my account of knowledge, it follows, roughly, that every way in which you could come to believe *p* is *either* unreliable as a way of getting to the truth, *or* such that it is not rational for you to regard it as a reliable way of getting to the truth. Suppose that you judged that this was the case. Then you would be committed to believing, of every way in which you could come to believe *p*, that *either* it is unreliable as a way of getting to the truth, *or* it is not rational for you to regard it as a reliable way of getting to the truth. But in that case you would be committed to believing, of every way in which you could come to believe *p*, that it is not rational for you to regard it as a reliable way of getting to the truth. It cannot be rational for you to believe, of some way in which you could come to believe *p*, that it is not in fact a reliable way of getting to the truth, but it is rational for you to believe that it is a reliable way of getting to the truth. That would be to believe something of the form '*not-q*, but it is rational for me to believe *q*'. But if it is rational for you to believe *q*, it cannot simultaneously be rational for you to believe *not-q*. So, if you believe the conjunction '*not-q* and it is rational for me to believe *q*', it cannot be the case that the first conjunct is rationally believed and the second conjunct true; if you are rational, you are in a position to appreciate this point. Thus, in the case that we are considering, you are committed to believing, of every way in which you could come to believe *p*, that it is not rational for you to believe it to be a reliable way of getting to the truth.

According to my account of rational belief, if some way of coming to believe *p* is such that it is not rational for you to believe it to be a reliable way of getting to the truth, then it is not a rational way for you to come to believe *p*. So, in the case that we are considering, you are committed to believing, of

every way in which you could come to believe  $p$ , that it is not a rational way for you to come to believe  $p$ . As I argued in §1, the judgment that it is not rational for you to believe  $p$  commits you to not believing  $p$ . This makes it plausible that if you are committed to believing, of every way in which you could come to believe  $p$ , that it is not a rational way for you to come to believe  $p$ , you are also committed to not believing  $p$ . So, we may conclude that it is necessary that, whenever you are not in a position to know  $p$ , there is a true judgment—roughly, ‘Every way in which I could come to believe  $p$  is either unreliable or at least it is not rational for me to regard it as reliable’—which rationally commits one to not believing  $p$ .

This explanation relies on my formulation of the norm of rational belief, which as I tried to make plausible in §4, is itself explained by the fundamental norm of correct belief. In this way then, it is plausible that the norm according to which, for every proposition  $p$  that one actually considers, one should believe  $p$  if and only if  $p$  is true does indeed explain the norms of rational belief and of knowledge. In this sense, this norm is indeed the fundamental epistemic norm. When interpreted in this way, that suggestive but metaphorical slogan, “Belief aims at the truth”, expresses a fundamental truth about belief.

## Notes

1. An earlier version of this paper was presented to audiences at Princeton University and at the Ohio State University. I am profoundly grateful to both audiences, and also to two different discussion groups of MIT and Harvard philosophy faculty and graduate students, and to Timothy Williamson, David Velleman, Alan Millar, Thomas Kelly and Alexander Bird, for extremely helpful comments.
2. Many other interpretations of this claim that “beliefs aim at the truth” have been proposed; see e.g. Velleman (2000, Essay 11). In fact, I suspect that on all the other interpretations that philosophers have proposed, the claim is either not true at all, or at least not an essential or constitutive truth about belief. But I shall not try to defend this suspicion here.
3. For more on my approach to normative concepts, see Wedgwood (2001).
4. Thus, it is quite possible for one’s engaging in a practice to “commit” one to something, even if it is in fact irrational for one to engage in (or to make any moves within) that practice at all; on this point, compare Broome (1999).
5. In my view, the term ‘ought’ is much more general than terms like ‘blameworthiness’: there are many cases in which someone does something that he ought not to do (or believes something that he ought not to believe) but is not “blameworthy” in any way. Similarly, ‘ought’ is much more general than the concepts typically expressed by ‘duty’ and ‘obligation’. There are many cases where you ought to do something (say, because it is necessary for your own happiness) but you have no “duty” or “obligation” to do it. Thus, my belief that the fundamental concepts of epistemic appraisal are normative concepts does not commit me to the view that Plantinga (1993, 11–29) stigmatizes as “epistemic deontologism”.
6. Objection: Why doesn’t this sufficient condition for normativity entail the controversial conclusion that the concept ‘true’ is normative? After all, the judgment ‘My

believing  $p$  would not be my believing a true proposition' commits one to not believing  $p$ . Reply: This judgment does indeed commit one to not believing  $p$ ; but that is only because its content entails ' $p$  is not true'. The fact that this judgment refers to *belief* has nothing to do with why it commits one to not believing  $p$ . Any judgment of the form ' $x$ 's  $\phi$ -ing  $p$  would not be  $x$ 's  $\phi$ -ing a true proposition' commits one to not believing  $p$ . So the fact that the judgment 'My believing  $p$  would not be my believing a true proposition' commits one to not believing  $p$  is not a *constitutive* feature of the concept 'true'.

7. A parallel point applies to practical reasoning, I believe. Suppose you judge that choosing to do  $x$  would be a correct or appropriate choice for you to make, while choosing to do  $y$  would not be a correct choice for you to make. Then you are committed to choosing  $x$  rather than  $y$ , if you choose either—and at all events, to not choosing  $y$ .
8. Indeed, it may be that the concept of “having disastrous consequences” only counts as normative for the practice of practical reasoning because if something  $x$  would have disastrous consequences, then any course of action  $y$  that prevents  $x$  from occurring would help to avert disastrous consequences, which counts as a reason in favour of doing  $y$ . Thus, it may be that the concepts that are most fundamentally normative for practical reasoning all apply to courses of action (or perhaps to choices or intentions). Concepts that apply to states of affairs more generally only count as normative for practical reasoning because judgments involving these concepts entail judgments that apply to courses of action (or to choices or intentions).
9. For a defence of this claim, see Wedgwood (1999b).
10. For criticisms of some of the leading attempts to individuate mental states in non-normative terms, see for example Bealer (1984). Explicitly normative accounts of the various types of mental state have been attempted by Morris (1992) and Brandom (1994) (although the details of these two accounts are in my opinion far from satisfactory).
11. I have something roughly akin to Kit Fine's (1994) conception of “essence” in mind here. For the distinction between “consequential” and “constitutive” senses of ‘essence’, see Fine (1995).
12. For a ground-breaking approach to partial beliefs based on this idea, see Joyce (1998).
13. What is it to “disbelieve”  $p$ ? On one view, the state of disbelieving  $p$  is just identical to the state of believing the negation of  $p$ . On another view, these two states are not identical since it is possible to disbelieve  $p$  even if one lacks the concept of negation. This second view is favoured by those, such as Rumfitt (2000), who wish to explain the concept of negation in terms of the principle that disbelieving  $p$  commits one to believing the negation of  $p$ , and vice versa. I shall remain neutral between these two views here.
14. I should emphasize that this norm only applies to beliefs (or doxastic or credal states). There are other attitudes that resemble beliefs in certain respects to which this norm does not apply. For example, suppose that one decides to treat a proposition as if true, for the purposes of making a certain decision. As Bratman (1993) has convincingly argued, this attitude need not be in any way incorrect even if the proposition is false. (Indeed, this attitude need not be irrational even if the proposition is *known* to be false.)
15. Since this norm only concerns propositions that one actually consciously considers, it does not support the suggestion of Alston (1989, 84), that “our central cognitive

aim is to amass a large body of beliefs with a favourable truth-falsity ratio". For a trenchant critique of the idea that we have any such "general cognitive aim", see Kelly (forthcoming).

16. I speculate that this norm also does not compare the state of suspending judgment about *p* with the state of having a mere "partial" degree of belief in *p*: from an epistemic point of view, a partial degree of belief is neither better nor worse than suspension of judgment. (There may of course be decisive *practical* considerations in favour of one or the other.)
17. This point is familiar from James (1979, 24–25).
18. Isn't one *rationaly required* (not merely *permitted*) to form this belief when in this condition? Yes, but we do not need to add that this rule requires forming this belief in this condition. We need to add that there is no rule that permits one to disbelieve *p*, or to suspend judgment about *p*, when one has an experience as of *p*'s being the case (and no special reason to think that one's experiences are unreliable in the circumstances). These rules are totalitarian: everything that these rules do not explicitly permit is forbidden.
19. As I shall use the phrase, the notion of "following" (as opposed to merely "conforming to") the rule is equivalent to what Millar (1991, 57–63) calls "exercising one's competence" with respect to the relevant "pattern of inference" or "rational connection".
20. As Goldman (1986, 44–48) would put it, you should aim to follow rules that are "locally reliable", not merely "globally reliable". For simplicity, let us interpret "the circumstances" as consisting in the region of space and time in which one considers the proposition in question. (Obviously, it is vague how extensive this space-time region needs to be. But there is a corresponding vagueness in the notion of "reliability", as we shall see.)
21. Here I am following Peacocke (1986, chap. 9), who also appeals to the truth-value of the beliefs that the "method" yields in "nearby possible worlds". However, there are a couple of differences between Peacocke's account and mine; see note 22 below.
22. There are a couple of differences between my conception of what it is for a set of rules to be "sufficiently reliable", and the necessary condition on knowledge that is proposed by Peacocke (1986, 134). Peacocke's requirement is that every "proper and improper initial submethod of one's total method" should yield a true belief in every nearby possible world in which it is used. I split this requirement into two: the requirement that the "set of rules" that one follows be "reliably error-avoiding" and the requirement that one's "method" be "reliably belief-yielding". My first requirement is stronger than the corresponding consequence of Peacocke's condition, while my second requirement is weaker.

My second requirement is weaker because I require only that one's total method should yield a belief in the proposition in question in every nearby world in which that method is used *and that proposition is true*. It surely casts no doubt on one's method of reaching belief in *p* if in some nearby world in which *p* is false, one uses this method, but this method does not yield any belief at all. (Although I only explicitly require that the "total method" be belief-yielding, this entails that the method's "proper initial submethods" will also be belief-yielding. If one's method reliably yields belief in *p*, then this method must also reliably yield belief in all the lemmas, or intermediate conclusions, that one must believe if that method is to yield belief in *p*.)

My first requirement is stronger because I require that the whole set of rules should be reliably error-avoiding, not just the method's "proper and improper initial submethods". Surely one's method for coming to believe  $p$  could be unreliable because, in the circumstances, some of the method's constituent rules could easily yield belief in a false proposition if they were followed in a slightly different *order* from the order in which they were actually followed.

23. Compare Harman's (1986, 44) suggestion that it is an "implicit commitment" of one's coming to believe a proposition on a certain basis that that basis is a "reliable indicator" of the truth of that proposition.
24. Indeed, someone might object that this conclusion rests on a sort of *level confusion*—a confusion between what could justify a higher-order attitude *about* a certain cognitive practice and what could justify that practice itself (compare Alston 1989, Essay 6).
25. Indeed, for most rules, one cannot distinguish the exact structure of the rule that one is following from that of various similar but subtly different rules. For example, one typically cannot know that one is following a rule  $R_1$  that permits one to judge  $p$  in response to a certain range of experiences  $E_1$ , as opposed to a very subtly different rule  $R_2$  that permits one to judge  $p$  in response to a very slightly different range of experiences  $E_2$ . For an argument for this point, see Williamson (2000, chap. 4).
26. Some epistemologists may insist that it is not rational to follow a rule unless one is in a position to grasp roughly what sort of rule one is following. I need not take a stand on this issue. But in fact I am inclined to disagree. Perhaps it is irrational to follow a rule if one actually tries to work out what sort of rule one is following, and is unable to do so; but it can still be rational to follow the rule in cases in which one makes no attempt to work out what sort of rule it is.
27. Exactly which rules, if any, are "basic" in this sense is one of the great questions of epistemology, which I cannot attempt to answer here. Very roughly, the view that I have defended elsewhere (1999a) is that the basic rational rules are the rules that play a certain sort of "pivotal" role in one's cognitive economy.
28. This proposal about rational belief may also be able to explain why rules that it is generally rational to follow must allow a "margin for error" in the sense that has been discussed by Williamson (2000, chap. 5).
29. In fact, this point suggests a promising approach to defeasibility in general. Whenever the support that one's being in a certain condition gives to a certain belief is *defeated*, this is always because it is no longer rational for one to believe that it is sufficiently reliable to form that belief in response to being in that condition. As has often been pointed out, for the justification for believing a proposition  $p$  to be defeated, it is neither necessary nor sufficient that there be evidence that supports believing the *negation* of  $p$ . It is not necessary because the defeating evidence might merely require suspension of judgment about  $p$ , not believing the negation of  $p$ . It is not sufficient since the evidence that supports believing the negation of  $p$  might itself be defeated by reasons for believing  $p$ . It is also not necessary that the defeating evidence make it rational to believe that this way of coming to believe  $p$  is *unreliable* in the circumstances. The defeating evidence might merely make it irrational for one not to *suspend judgment* about whether this way of coming to believe  $p$  is reliable in the circumstances.
30. *These* rules will certainly not be rules that it is "generally rational" to follow: even if it is rational to follow these rules in some cases, it will not be rational to continue

- following these rules if one acquires sufficiently strong evidence against the corresponding background belief.
31. This principle is not itself an epistemic norm. Admittedly, it refers to what it is “rational for one to believe”, and so presupposes that there *are* norms of rational belief. This may be avoidable in a fuller account. Even if it is not avoidable, however, this does not vitiate the explanation, since this principle does not itself say anything about the *content* of the norms of rational belief.
  32. Even though I claim that reliability-in-the-circumstances is a necessary condition for knowledge, there are at least two ways in which I depart from the more familiar forms of reliabilism, such as the form advocated by Goldman (1979). First, my approach is compatible with an *internalist* conception of rationality. Rationality, on my account, requires only that it must be *rational for one to believe* the rules that one is following to be reliable, not that these rules must actually be reliable. The rationality of believing these rules to be reliable may be determined purely by “internal” facts about one’s mental states—not by facts about the external world that could vary while those internal facts remained unchanged. Second, according to my account, one can only come to know something by following rules that it is rational for one to follow. If one forms a belief in any way other than by following such rational rules, that does not count as the acquisition of knowledge, and indeed, there is no need to determine whether the belief is formed as a result of a sufficiently reliable “process” or not.
  33. Thus, my view respects the point, which was made independently by both Zagzebski (1994) and Merricks (1995), that any conception of “warrant” according to which warrant does not imply truth will be vulnerable to a Gettier-style counter-example.
  34. This example was due to Carl Ginet, but was first mentioned in print by Goldman (1976). In dealing with this example, I am closely following Peacocke (1986, 136–37). I can therefore endorse Peacocke’s reply to the strengthened barn façade example that Saul Kripke put forward in his unpublished 1985 Gareth Evans Memorial Lecture.
  35. These two cases were suggested to me as counterexamples to the “no false lemmas” requirement—the first by Timothy Williamson, the second by Gideon Rosen.
  36. Why aren’t there Gettier cases where a rational belief fails to count as knowledge because some of the rules that one followed in holding that belief, though reliably error-avoiding, were *unnecessarily restrictive*? There seem not to be any such cases; and this is a puzzle for my conception of knowledge. I think the answer must be that no rule that is in the relevant sense “unnecessarily restrictive” can be a rational rule to follow. A rule  $R_1$  counts as unnecessarily restrictive just in case there is some other rule  $R_2$  available to one such that it is rational to regard both  $R_1$  and  $R_2$  as sufficiently and equally reliable, where  $R_1$  and  $R_2$  differ from each other only in that  $R_2$  is less restrictive than  $R_1$ . Suppose that for  $R_2$  to be “available to one” one must actually have some *mastery* of  $R_2$ . Then if one formed a belief through following rule  $R_1$ , and not through following  $R_2$ , one would essentially be treating certain evidence as relevant to justifying the belief in question, even though the evidence is in fact irrelevant; and that, I think, would be irrational. (Cases where one follows an excessively restrictive rule like  $R_1$  must be distinguished from cases in which one’s belief is over-determined, formed through following *both*  $R_2$  and some other rule  $R_3$ .)
  37. Some philosophers, such as Nelkin (2000), claim that it is not rational for you to believe that I won’t win the lottery. But even these philosophers regard this claim

as prima facie implausible and in need of special defence. There is also reason to suspect that this claim will have sweepingly sceptical consequences (especially with respect to our knowledge of the future).

38. Admittedly, contexts in which it is true to say that you know that I won't win the lottery are rather odd. One contextualist who gives a subtle explanation of why this is so is Craig (1990, 98–103). Although I endorse a contextualist account of the lottery, I do not accept contextualist solutions to the problem of scepticism; but I cannot discuss these issues here.
39. For example, Craig (1990) argues that knowledge makes one a valuable source of information for others in a way in which mere true belief does not; and Williamson (2000, 75–89) argues that knowledge plays a more robust or resilient role in explaining action than mere true belief.

## References

- Alston, W. 1989. *Epistemic Justification* (Ithaca, New York: Cornell University Press).
- Audi, R. 1993. *The Structure of Justification* (Cambridge: Cambridge University Press).
- Bealer, G. 1984. "Mind and Anti-Mind: Why Thinking Has No Functional Definition", *Midwest Studies in Philosophy* 9, 283–328.
- Broome, J. 1999. "Normative Requirements", *Ratio* 12, 398–419.
- BonJour, L. 1980. "Externalist Theories of Empirical Knowledge", *Midwest Studies in Philosophy* 5, 53–74.
- Brandom, R. 1994. *Making It Explicit* (Cambridge, MA: Harvard University Press).
- Bratman, M. 1993. "Practical Reasoning and Acceptance in a Context", *Mind* 101, 1–15.
- Conee, E., and Feldman, R. 1998. "The Generality Problem for Reliabilism", *Philosophical Studies* 89, 1–29.
- Craig, E. 1990. *Knowledge and the State of Nature* (Oxford: Clarendon Press).
- Davidson, D. 1980. *Essays on Actions and Events* (Oxford: Clarendon Press).
- Fine, K. 1994. "Essence and Modality", *Philosophical Perspectives* 8, 1–16.
- Fine, K. 1995. "Senses of Essence", in W. Sinnott-Armstrong (ed), *Modality, Morality, and Belief* (Cambridge: Cambridge University Press).
- Gettier, E. 1963. "Is Justified True Belief Knowledge?", *Analysis* 23, 121–23.
- Goldman, A. 1976. "Discrimination and Perceptual Knowledge", *Journal of Philosophy* 73, 771–91.
- Goldman, A. 1979. "What is Justified Belief?", in G. Pappas (ed.), *Justification and Knowledge: New Studies in Epistemology* (Dordrecht: Reidel).
- Goldman, A. 1986. *Epistemology and Cognition* (Cambridge, MA: Harvard University Press).
- Harman, G. 1973. *Thought* (Princeton, NJ: Princeton University Press).
- Harman, G. 1986. *Change in View* (Cambridge, MA: MIT Press).
- James, W. 1979. *The Will to Believe and Other Essays in Popular Philosophy* (Cambridge, MA: Harvard University Press).
- Joyce, J. M. 1998. "A Nonpragmatic Vindication of Probabilism", *Philosophy of Science* 65, 575–603.
- Kelly, T. P. forthcoming. "Epistemic Rationality as Instrumental Rationality: A Critique", *Philosophy and Phenomenological Research*.
- Lewis, D. K. 1996. "Elusive Knowledge", *Australasian Journal of Philosophy* 74, 549–67.
- Merricks, T. 1995. "Warrant Entails Truth", *Philosophy and Phenomenological Research* 55, 841–55.
- Millar, A. 1991. *Reasons and Experience* (Oxford: Clarendon Press).
- Morris, M. 1992. *The Good and the True* (Oxford: Clarendon Press).
- Nelkin, D. K. 2000. "The Lottery Paradox, Knowledge, and Rationality", *Philosophical Review* 109, 373–409.
- Nozick, R. 1981. *Philosophical Explanations* (Cambridge, Mass.: Harvard University Press).
- Peacocke, C. 1986. *Thoughts: An Essay on Content* (Oxford: Blackwell).
- Peacocke, C. 1999. *Being Known* (Oxford: Clarendon Press).

- Plantinga, A. 1993. *Warrant: The Current Debate* (New York: Oxford University Press).
- Rumfitt, I. 2000. "'Yes' and 'No'", *Mind* 109, 781–823.
- Velleman, J. D. 2000. *The Possibility of Practical Reason* (Oxford: Clarendon Press).
- Wedgwood, R. 1999a. "The A Priori Rules of Rationality", *Philosophy and Phenomenological Research* 59, 113–131.
- Wedgwood, R. 1999b. "The Price of Non-Reductive Moral Realism", *Ethical Theory and Moral Practice* 2, 199–215.
- Wedgwood, R. 2001. "Conceptual Role Semantics for Moral Terms", *Philosophical Review* 110, 1–30.
- Williams, B. 1978. *Descartes: The Project of Pure Enquiry* (Harmondsworth: Penguin Books).
- Williamson, T. 2000. *Knowledge and its Limits* (Oxford: Clarendon Press).
- Zagzebski, L. 1994. "The Inescapability of Gettier Problems", *Philosophical Quarterly* 44, 65–73.