1

Is Rationality Normative?

1.1 The word ‘rational’

The word ‘rational’ – and its equivalents in other languages – run through the texts of Western philosophy like the so-called “rogue’s yarn” that ran through every piece of rope that belonged to the 18th-century British navy. The word ‘rational’ and its cognates are unavoidable in translations of Plato’s Republic, Aristotle’s Ethics, Aquinas’s Summa Theologica, and Kant’s First and Second Critique – and most of the other central texts of Western philosophy as well.

In many of its appearances in these philosophical discussions, the word’s primary reference is to the faculty of reason – often conceived as it is by Plato (Republic 439c-d) as the part of the soul that is involved in reasoning and calculation. The word is arguably also used in this way when Aristotle (Nicomachean Ethics 1098a3-12) defined human beings as essentially “rational animals” – that is, animals who possess this special faculty. Again, when Aquinas (Summa Theologica Ia–IIae, 66.4) defines the will as the “rational appetite” (appetitus rationalis), what he seems to mean is that the will is the form of appetite or desire that involves the deployment of this distinctive faculty.

The word ‘rational’ came into English from these ancient and medieval Latin texts. As soon as it appears in English, the word’s primary reference seems again to be to this special faculty. So, for example, Hobbes and Locke repeatedly refer to human beings as “rational creatures”, because we are creatures who have this special faculty. Starting at least as early as Locke, however, we find another use, in which the word is used to refer to actions, choices, and beliefs (and the like) that are in some sense based on or derived from the proper use of the faculty of reason.¹

This last use of the term is a normative use. To use our faculty of reason “properly” is to use it as it should be used. When a belief or a choice derives from the proper use of this faculty, it thereby has what in the broadest sense could be called a kind of virtue or excellence – a

¹ See, for example, Locke’s Essay concerning human Understanding, I.ii.25, and I.iv.25.
feature in virtue of which beliefs (and the like) can count as good, or worthy of a certain kind of commendation.

There can, I think, be no doubt that the word ‘rational’ sometimes has this normative meaning. It is common, after all, for words that refer to a particular sort of activity or faculty to come to refer to instances of that activity or uses of that faculty that manifest the distinctive excellence that is associated with that activity or faculty. Words like ‘artist’ can be used both in a value-neutral way (as when one might damn someone by calling them a “talentless artist”), and also in a more normatively loaded way (as when one might express one’s admiration for a person by simply calling them an “artist”).

Our linguistic intuitions seem to support the conclusion that the word ‘rational’ can express a normative concept of this sort. Suppose that you are reasoning aloud with a friend, and your friend turns to you and says, “You are being irrational.” You will surely take this a criticism — indeed, most likely as a harsh and stinging criticism at that.² This seems to be because it is widely assumed that the term ‘irrational’ stands for a kind of failing or shortcoming: if your thinking is irrational, then there is something wrong with it, something bad or defective about it. In other words, it is assumed that irrationality is a kind of defect or flaw or blemish, and rationality is the corresponding kind of virtue or flawlessness.

This point does not yet establish that the concept expressed by ‘rational’ is a normative concept. You would also take it as a stinging criticism if your friend said, “You’re fat”, or if your friend said, “You have the mental capacities of a slug.” This does not show that the concepts of being fat and of having the mental capacities of a slug are normative concepts. However, it is plausible that there is a difference between the two cases. It seems plausible that irrationality is always and necessarily some kind of defect or flaw or blemish — whereas being fat is not always a defect (it is not a defect in Sumo wrestlers), and having the mental capacities of a slug is not necessarily a defect (it would not be a defect in a world in which slugs were super-intelligent angelic beings). This makes it plausible that it is a conceptual truth that irrationality is some kind of defect or flaw — that is, in effect, that the concept of “irrationality” is itself a normative concept.

There are also other linguistic intuitions that support the same conclusion. Some of these other linguistic intuitions have to do with the connection between the word ‘rational’ and words like ‘justified’. It seems that there is a way of using the terms ‘rational’ and ‘justified’

² Compare Hussain (2007, 29).
so that phrases like ‘rational belief’ mean exactly the same thing as ‘justified belief’. (This is the sense of the terms ‘rational belief’ and ‘justified belief’ that is of particular interest to epistemologists.) In this book, I shall be using the terms ‘rational’ and ‘justified’ in this way. But the term ‘justified’ surely expresses a normative concept. There are many ways in which things can be correctly called “justified”: for example, there is both moral and legal justification, as well as the kind of justification of belief that is of special interest to epistemologists. Indeed, the etymology of the word ‘justified’ seems strongly to suggest that the word is likely to express a normative concept: ‘justified’ is obviously cognate with ‘just’, and with the Latin word ‘jus’, which means law (not in the sense of an explicit regulation like a statute, but in the sense of the whole body of rules or principles that ought to be used by public authorities to order human society).

It would be most implausible to suggest that the word ‘justified’ is ambiguous in the same way as words like ‘bank’ and ‘palm’. In contemporary English, there are two completely unrelated concepts expressed by ‘bank’ (river bank and money bank), each of which must be learned separately as a different meaning of ‘bank’. This makes it plausible to say that there are in fact two different words that can be pronounced ‘bank’. The situation with ‘justified’ seems completely different. When we talk about “morally justified behaviour” and about “justified beliefs”, we are surely using the word in closely related senses. In general, it seems plausible that all the concepts expressed by the term ‘justified’ are sufficiently closely related that they all count as normative concepts. So, when ‘rational’ is used in a way that makes ‘rational belief’ equivalent to ‘justified belief’, the term ‘rational’ also expresses a normative concept.

The terms that are especially strongly associated with normative concepts include the deontic modal terms like ‘ought’ and ‘should’. It seems plausible that the statement that it is (in this normative sense) “irrational” for you to have a certain belief or to make a certain choice entails that (in some sense) you “ought not” to have that belief or make that choice; to say that it is “rational” for you to have this belief or to make this choice entails that (in the relevant sense) that it is not the case that you ought not to have the belief or make the choice — which seems to be equivalent to saying that you may permissibly have that belief or make that choice. ³ The same point holds of terms like ‘justified belief’. To say that you are justified in believing \( p \) entails that it is, in a certain sense, permissible or all right for you to

³ Technically, this is known as the thesis that ‘ought’ and ‘may’ are duals of each other (just like ‘must’ and can’, or ‘all’ and ‘some’). For an argument for this thesis, see Wedgwood (2007a, Section 5.1).
believe \( p \); to say that you are not justified in believing \( p \) entails that it is, in a corresponding sense, impermissible for you to believe \( p \) — or other words, that you ought not to believe \( p \).

Admittedly, this is, so far, only a very rough indication of what it means to say that the concept of “rationality” is a normative concept. A more precise account will be given in the next chapter. For the rest of this chapter, I shall assume, in order to fix ideas, that if the concept of “rationality” really is normative, then the following two principles hold: (a) rational mental states (like beliefs) and rational mental events (like events in which one forms or revises one’s beliefs) are, in a way, good or worthy of commendation; (b) if a thinker is rationally required to \( \phi \), there is a sense in which the thinker ought to \( \phi \).

In fact, it seems to me that it should not be controversial that the English word ‘rational’ can be used to express a normative concept of this kind. What is controversial is whether the term is used in this way when it figures in the most influential contemporary theories of rationality that have been developed by philosophers, and by some other formal theorists like economists and statisticians.

1.2 Formal theories of rationality

The word ‘rational’ is often used in certain formal theories, many of which have been developed with great mathematical sophistication. The most striking examples of these formal theories are the theories of rational decision that were developed by such theorists as John von Neumann and Oskar Morgenstern (1944), who in turn were building on the pioneering work of F. P. Ramsey (1926). These formal theories of rational decision were integrated with formal theories of rational degrees of belief by L. J. Savage (1954), who was in turn building on the work of Bruno de Finetti (1937). This field of research has made considerable progress over the years, with the development of many different formal theories of rational degrees of belief and of rational decision; the most prominent recent contributors include such theorists as Richard Jeffrey (1983 and 2004) and James M. Joyce (1999).

In addition to these theories of rational decision and rational degrees of belief, some formal theories have been developed that treat belief, not as a phenomenon that comes in degrees, but simply as an on/off phenomenon — so that the only possible form of belief that one can have towards a proposition is simply believing it, and the only alternative to simply believing the proposition is simply not believing it. Within a framework that treats belief in this way, the best example of a formal theory of rationality is probably the logical belief revision theory of Carlos Alchourrón, Peter Gärdenfors, and David Makinson (1985).
All these formal theories have a fundamental similarity to each other. Each theory describes a kind of coherence that can hold among the thinker’s beliefs, intentions, and other mental states (such as preferences, experiences, memories, and the like) at a particular time. This kind of coherence depends purely on the nature and content of the mental states that are present in the thinker’s mind at the time in question. Intuitively, the kind of coherence in question is purely a matter of these mental states harmonizing or fitting together properly, as opposed to being in some kind of conflict or discord with each other.

Typically, this sort of coherence requires at least a kind of logical consistency among the beliefs and intentions that the thinker has at the relevant time (or at least among the beliefs that the thinker holds with certainty at that time). For example, according to many theories, coherence requires that the thinker’s degrees of belief must be probabilistically coherent: that is, it must be possible to represent these degrees of belief by means of a probability function. Many theories also insist that to be coherent, the thinker’s preferences must meet a number of conditions (such as transitivity, and independence or the “sure-thing principle”) which make it possible to represent those preferences by means of a utility function. Proponents of these formal theories typically say that at every time, the system of mental states (such as the beliefs and intentions and the like) that a thinker has at that time is rational only if it forms a coherent system in this way. Since these conditions of coherence apply only to the system of mental states that the thinker has at a single time, they could be called conditions of synchronic coherence.

Most of these formal theories also describe a way in which a thinker might over time revise her mental states (such as her beliefs and intentions and the like) that is in a sense “conservative”. With each of these revisions, it is assumed that the thinker must revise her old system of mental states so that the new system conforms to a certain constraint (for example, the new system must include an attitude of complete certainty towards certain propositions that the agent has just learnt at the time in question). For one of these revisions to be “conservative” in the relevant way, the thinker must, according to some measure, conserve as much as possible of her past system of mental states while switching over to a new system that (a) meets the relevant constraint and (b) meets all the conditions of synchronic coherence that we have just discussed.

For example, one of the most famous theories of this sort says that whenever a thinker with a probabilistically coherent system of degrees of belief learns some new fact with certainty, she must switch to the system of degrees of belief that corresponds to the result of conditionalizing her old system on that new fact. This is equivalent to saying that her new
system must meet the constraint (it must involve believing this new fact with certainty); it must meet the conditions of synchronic coherence (above all, it must also be probabilistically coherent); and it must conserve (or “hold fixed”) the conditional probabilities of all propositions, conditional on the new fact that has just been learnt.

In effect, this sort of conservatism in revising one’s beliefs and intentions is a kind of coherence over time. The new system must harmonize or fit together with the old system, together with the new constraint that the new system has to meet. If the thinker were to revise her system of mental states in a less conservative fashion – for example, if she were to undergo a radical conversion, as in the Biblical story of Paul on the road to Damascus (Acts 9: 3–9) – her new system of mental states would lack the kind of coherence or continuity with the old system that makes it possible (according to these formal theories) to view this change to the thinker’s mental states as a process of rational reasoning.

In this way, then, these formal theories of rationality define these sophisticated notions of coherence, including both synchronic coherence (coherence at a time) and diachronic coherence (coherence over time). These formal theories of rationality play a large role in economics, game theory, certain branches of political science, and other forms of social theory. This is because in these theories, it is assumed that the behaviour being studied is more-or-less rational at least most of the time: so finding out what will happen if people behave rationally is at least a rough guide to what will actually happen.

The question that I want to focus on in the rest of this chapter is this: Do these formal theories of rationality use the term ‘rational’ to express a normative concept, of the kind that I have discussed in Section 1.1? Or do they use the term in a different sense – for example, to express a special technical concept of some kind?

1.3 Problems with non-normative interpretations of ‘rational’

If ‘rational’ (as it is used in these formal theories) is not a normative concept, but a special technical concept, then what sort of concept is it? Some philosophers will propose that this concept is really a logical concept, like the concept of logical consistency or the like, rather than a genuine normative concept.4

It is certainly true that in a sense, these formal theories of rationality aim to give a “definition” of rationality in broadly logical terms: for example, the subjective Bayesians and the expected utility theorists set out to say what it is for a set of preferences and

---

4 For clear statements of these proposals, see especially Horwich (1998b) and Schroeder (2003).
credences to be rational in terms of various formal conditions of coherence, which could be regarded as broadly logical in character. On this proposal, then, each of these formal theories is simply stipulating a definition of some kind of coherence, by using these logical and mathematical devices, and then uses the term ‘rational’ as a shorthand to express the concept of this sort of coherence.

However, this proposal faces a number of problems. First, there are many kinds of coherence that could be defined by means of logical and mathematical devices – why is there anything interesting about this particular kind? There is no point in simply stipulating definitions of concepts for its own sake. There are infinitely many stipulative definitions that could be given. Stipulating a definition of a concept has a point only if it is clear what theoretical role the stipulated concept is going to play. If the theory in question is an empirical theory, then we will need some justification for thinking that the concept can play this role within an empirical theory. If the theory in question is a purely mathematical theory, then we need some justification for thinking that there is some interest in developing a mathematical theory in which the concept plays that role.

However, this interpretation of these formal theories of rationality as simply stipulating definitions of coherence makes it quite unclear what justification there is for these stipulations. On the one hand, if the chief purpose of these stipulations is to play a role in empirical theories like the forms of economics and social theory that we have just discussed, then it is quite unclear what justification we have for thinking that by and large, most actual agents will approximate to satisfying these conditions of coherence in their beliefs and choices. Why should it be reasonable for economists and social scientists to assume that most of the time, the behaviour that they study exhibits this particular kind of coherence (rather than some other kind of coherence that could be defined)?

If, on the other hand, this stipulated notion of coherence is designed only to play a role within purely mathematical theory, then it is unclear what interest there is in the mathematical theory that is concerned with one of these notions of coherence. Indeed, every one of these notions of coherence is mathematically just as interesting as every other. There is no reason to favour any one definition over any other.

Indeed, in general, if ‘rational’ just expressed a purely logical concept, these different formal theories are simply stipulating definitions of different kinds of coherence, and not disagreeing with each other about exactly which processes of reasoning count as “rational”. But this is precisely not what we think of these different theories of rationality. We intuitively think of these different theories as disagreeing with each other. That is, we intuitively suppose that there is a single concept of ‘rationality’, and these different
theories are offering rival accounts of what it is for a system of mental states to count as rational. But if this is what these theories are doing, then we cannot regard these theories as simply stipulating definitions of some notion of coherence.

Finally, we might wonder about why these formal theories of rationality are so often put forward as though they had some philosophical importance. It seems that they are thought to have philosophical importance because they were meant to count as contributions to the same debates as more traditional discussions of “justified belief” or “inductively confirmed theories” and the like. But if the formal theories were simply stipulating definitions of formal varieties of coherence, and using the term ‘rational’ as shorthand for those varieties of coherence, there would be no reason to think that these formal theories had any relevance whatsoever to those traditional debates.

Thus, it seems more plausible that the “definitions” given by these formal theories of rationality are attempts at giving a substantive theory of rationality, in much the same way as consequentialists and their opponents in ethics give substantive theories of moral rightness and wrongness. It is not plausible to require that the correct substantive theory of rightness and wrongness must be known by everyone who possesses the concepts of “rightness” and “wrongness”. In a similar way, the correct substantive theory of rationality also need not be known by everyone who possesses the concept of “rationality” (even if this account is ultimately accessible by means of a sort of a priori reflection to anyone who possesses this concept). Even while philosophers dispute about what the correct substantive theory of rationality is, they all continue to use the very same concept of “rationality”. So possessing this concept does not consist simply in knowing the correct definition; a different sort of account of what it is to possess this concept is required.

For these reasons, then, this purely logical interpretation of the concept of rationality faces severe problems. By contrast, as I shall now argue, if ‘rational’ expresses a normative concept, these problems do not arise.

When the term ‘rational’ expresses a normative concept, “rationality” is the proper use of our reasoning faculties: it is clear why it is interesting to find out what it is to use these faculties properly. The different formal theories give rival accounts of what it is to use these faculties properly; in this way, our intuitive sense that these theories disagree with each other about what it takes to be rational is simply correct. Moreover, our interest in these theories is not purely in the mathematical structure of the abstract forms of coherence that they define; we are interested in these theories as potentially giving the correct account of what it is for us to use our reasoning faculties properly (in the relevant sense).
Finally, if the term ‘rational’ expresses this normative concept, we may also be able to explain the role that this concept plays in economics and in other forms of social sciences. Perhaps we can assume that in certain domains, people do most of the time use their reasoning faculties properly – or at least that people come sufficiently close to using their reasoning faculties properly so that in aggregate the large-scale social outcome does not differ too much from what it would have been if everyone had been reasoning properly. Markets and economic behaviour may well be one of these domains. So it is not out of the question that this normative interpretation of the concept of “rationality” can be reconciled with the central role that the concept plays in economics and in other social sciences.

In this way, then, it seems to me that there are some powerful reasons for interpreting the term ‘rational’, as it appears in these formal theories, as expressing a normative concept of some kind. However, as we shall see in the rest of this chapter, there are also some serious objections that can be raised against this interpretation.

1.4 Four objections to the normative interpretation of ‘rational’

As I have explained above, I am assuming here that if the term ‘rational’ expresses a normative concept, then one conclusion that follows is that in some sense of the term, we always “ought” to be as we are rationally required to be. Yet there are several objections that can be raised against this conclusion. I shall focus on four such objections here.

(a) The strike of the demon

The first objection that I wish to consider has been clearly raised by Andrew Reisner (2011). This objection can in fact be raised against every conceivable requirement of rationality. Consider any example of a putative rational requirement – any condition that we are plausibly rationally required to meet. For example, we might consider the alleged requirement that we should not simultaneously be totally confident both of a proposition \( p \) and of its negation \( \neg p \). Let us concede that it would be irrational for you not to conform to this requirement; that is, in this example, you are rationally required not to be simultaneously totally confident both of \( p \) and of \( \neg p \).

However, we can easily imagine that in this case conforming to this alleged rational requirement will be utterly disastrous. For example, suppose that a demon will destroy the
world unless you are simultaneously totally confident both of \( p \) and of \( \neg p \). In this case, it may seem compelling that if your \( \varphi \)-ing would provoke the demon to destroy the world, then it cannot true that you ought to \( \varphi \). So, if avoiding this combination of attitudes will provoke the demon to destroy the world, it seems not to be true that you ought to avoid this combination of attitudes. That is, it seems not to be true that you ought not to be simultaneously totally confident both of \( p \) and of \( \neg p \). Or alternatively, we might suppose that an eccentric billionaire will use his resources to save huge numbers of people if you are simultaneously totally confident of both of \( p \) and of \( \neg p \). In this case, again, it seems not to be true that you ought not to be simultaneously confident both of \( p \) and of \( \neg p \).

If we assume that you are rationally required not to be simultaneously confident both of \( p \) and of \( \neg p \), this is one rational requirement of which it seems not to be true that you ought always to conform to it. So ‘is rationally required to’ does not always imply ‘ought’ – contrary to what I am assuming to follow from the claim that rationality is a normative concept. In this way, this argument could be used to raise doubts about the claim that rationality is a normative concept.

In principle, an alternative conclusion could be drawn from the argument. Specifically, it could be concluded that since rationality is a normative concept, it cannot in fact be true that rationality always requires us not to be simultaneously totally confident both of \( p \) and of \( \neg p \). In general, this style of argument could be used to argue that there are no forms of coherence that rationality always requires. If it is also assumed to be a conceptual truth about rationality that if there are any rational requirements, there must certain forms of coherence that are always required of us, then this argument could be used for a kind of “error theory” about rational requirements – that is, for the conclusion that there are no rational requirements.

At all events, Reisner (2011, §3) takes the argument to show that rational requirements are not normative, not that there are no rational requirements. He offers the following diagnosis of why this objection arises:

> The truth of any particular normative requirement is, like any other ought, determined by the various features of the world on which normativity is dependent.

---

5 This objection seems to be clearly modelled on the “strike of the demon” arguments that have been studied by Rabinowicz and Rønnow-Rasmussen (2004).

6 This “error theoretic” argument is modelled on the arguments against moral requirements that are due to Mackie (1977) and Richard Joyce (2001). Compare how Kolodny (2007a) understands his arguments against “coherence” theories of rationality.
Eccentric billionaire examples can be generated for any requirement of rationality that could be given as a normative requirement; sometimes it is the case that it would be so bad to be rational that one ought not to be.

Rationality, on the other hand, is dependent on some complex of the logical features of the relations amongst an agent’s mental states and their contents; the rational supervenes strictly on the mental, and this is not the case for the normative. That it would be extremely bad for an agent to have mental states that would qualify an agent as (locally) rational does not affect the truth of the matter about whether or not an agent’s mental states are such that she is (locally) rational.

According to this diagnosis of the objection, normative requirements are just fundamentally different phenomena from rational requirements. Rational requirements supervene on local features of the agent’s mental states, while normative requirements supervene on a much wider array of “features of the world”.

(b) Rational beliefs in false normative propositions

It seems that propositions that are rationally believed can sometimes be false. So suppose that you have a rational false belief about what you ought to do: you rationally believe that you ought now to φ, but in fact you are wrong; in fact, it is not the case that you ought to φ – instead, you ought now not to φ. In this case, it may well be irrational for you to give up this rational false belief. In that sense, you are rationally required now to believe that you ought now to φ – even though this belief is in fact false, and the truth of the matter is that you ought not now to φ.

At least under certain further assumptions, it seems that we can make it plausible that in this case, you are rationally required either now to φ, or at least to intend now to φ. First, let us assume that φ-ing is a mental state or a mental activity of the sort that can be rationally required of us in certain situations. Then, surely, if you are rationally required now to believe that you ought now to φ, then (given these assumptions about what φ-ing is), you must also be rationally required now to φ. So, by hypothesis, this is a case in which you are rationally required now to φ, but it is not true that you ought now to φ. So this seems to be a counterexample to the claim that rationality is a normative concept.

7 This problem is effectively a version of what is known as “Ewing’s problem”, after its presentation by Ewing (1953, 144f.).
Secondly, let us assume that \( \phi \)-ing is not a mental state or activity of the sort that can be directly rationally required of you, but is instead a type of action such that you know that you will now \( \phi \) if and only if you now form an intention to do so. Given that you are rationally required now to believe that you ought now to \( \phi \), and that you know that you will now \( \phi \) if and only if you form an intention now to \( \phi \), surely it is plausible that you are now rationally required to form an intention now to \( \phi \). So it seems that rationality requires you to form an intention now to \( \phi \). But if ‘rational’ were a normative concept, then it would have to be true, at least in some sense, that you “ought” to form the intention to \( \phi \). But, by hypothesis, the truth of the matter is that you ought not now to \( \phi \). Can it really be true that you ought to form an intention to \( \phi \), given that you ought not to \( \phi \)?

In short, it seems that rational normative beliefs, of the kind that one might express by a first-person present-tensed statement of the form ‘I ought now to \( \phi \)’, can be false. But when one has a rational belief of this sort, in a situation in which it is irrational for one to abandon the belief, this belief can easily create further rational requirements (such as a rational requirement to \( \phi \), or to intend to \( \phi \)); and these further rational requirements will come into conflict with the fact that in this case one ought not to \( \phi \) if rationality is indeed (as I have argued above) itself a normative concept.

In principle, some philosophers might defend the view that it is not even possible for anyone to have a rational false belief in any proposition that they might express by saying something of the form ‘I ought now to \( \phi \)’. This view is occasionally attributed to Joseph Butler, who is sometimes interpreted (quite wrongly, in my view) to have held that conscience is infallible.\(^8\) According to this view, propositions of this kind would have the remarkable feature that when it comes to these propositions, rational belief guarantees truth. But at least on the face of it, it seems only too easy to imagine cases in which I rationally believe the proposition that I would express by saying ‘I ought now to \( \phi \)’ even though the proposition is false. The simplest case is where I have received utterly compelling but in fact misleading evidence that I ought to \( \phi \) (for example, suppose that a sublime oracle that has the most amazing track record for reliability announces that I ought now to \( \phi \)), and I base my belief that I ought to \( \phi \) on this misleading evidence. In this case, my belief is both rational and false. Thus, it seems prima facie highly plausible that with these beliefs just as with almost all others, rationality does not guarantee truth.

\(^8\) For the attribution of this view to Butler, see Anscombe (1958). For an alternative interpretation of Butler that acquits him of holding this view, see Wedgwood (2007b, 188).
So it seems that it must be possible for beliefs that one could express by saying something of the form ‘I ought to φ’ to be both rational and false. As we have just seen, this point seems to create some serious difficulties for the view that rationality is a normative concept.

(c) ‘Ought’ implies ‘can’

I am assuming here that if rationality is a normative concept, then the claim that one is rationally required to φ implies that one in a sense “ought” to φ. But it is widely held that ‘ought’ implies ‘can’; and is it true that one can do whatever is rationally required of one?

There are at least two reasons for thinking that this may be a serious problem for normative interpretations of rational requirements. First, rational requirements typically apply to beliefs or other similar mental states like inferences and the like; and it seems that we cannot just form or revise our beliefs, or draw inferences from our premises, at will, in the way in which we can normally act at will.Œ If we can’t conform to these rational requirements at will, is it really true that we “can” conform to them at all (in the sense of ‘can’ that is relevant to the principle that ‘ought’ implies ‘can’)?

Secondly, some rational requirements seem to be so fantastically idealized that one might wonder whether it is possible in any sense for us to conform to them. Many formal theories of rationality assume that perfectly rational thinkers will have a set of preferences and degrees of belief that is complete, in the strong sense that it includes an attitude towards every proposition in an infinite propositional algebra (that is, a set of propositions that is closed under operations like negation and disjunction and the like). It may seem most doubtful whether it is possible for any finite creature like a human being to have a set of attitudes that is complete over an infinite domain of propositions of this sort.

Admittedly, we could amend these formal theories so that they do not require this sort of completeness. But since all these formal theories use the devices of logic and probability theory to characterize the relevant kind of coherence, it will be hard for them to avoid imposing highly idealized demands altogether. The point is particularly clear with probabilistic coherence. If your degrees of belief were probabilistically coherent, then you would never have any degree of belief in any logical truth that was anything short of the maximum degree of confidence; you would never have any degree of confidence in any

---

Œ For some particularly influential statements of this objection, see Alston (1989, 115-152) and Plantinga (1993, 37f.).
logical falsehood that was anything greater than total disbelief; and you would never have
a different level of confidence in any two propositions that were logically equivalent to
each other.

Even if this probabilistic theory does not require logical omniscience – that is, even if it
allows for some rational thinkers to have no attitude whatsoever towards some
propositions, even towards some logical truths – it still requires a kind of logical infallibility.
That is, it seems very hard to see how any probabilistic theory can allow any rational
thinker to have any degree of belief other than total confidence in any logical truth. We
might well wonder whether such logical infallibility is even possible for normal human
thinkers like you and me.

Suppose that rationality requires such logical infallibility, and that it is not actually
possible for us to be logically infallible in this way. Then if every rational requirement
entailed a corresponding ‘ought’, we would have a violation of the principle that ‘ought’
implies ‘can’. Given the plausibility of that principle, this seems to cast doubt on the view
that every rational requirement entails a corresponding ‘ought’. That is, it seems to cast
doubt on the view that rationality is normative.

(d) What’s good about coherence?

On several views of the matter, including the views that will be defended in this book,
rationality requires nothing more than some kind of coherence between the relevant agent’s
mental states and mental events – at least in the broad sense of the term ‘coherence’ that I
am using here.

However, as Niko Kolodny (2005)¹⁰ has asked, what is good about coherence? Isn’t it just a
mere pretty pattern of mental states and events?¹¹ Such coherence is compatible with
having beliefs in propositions that are false, or with having intentions that will lead one to
cause appalling disasters. In some cases, indeed, if one is sufficiently unfortunate, then
making one’s mental states more coherent may only spread the disastrous errors more
widely through one’s system of mental states.

¹⁰ See also other writings of Kolodny (2007, 2008a, 2008b). Other philosophers have sought to cast
doubt on certain specific alleged rational requirements by raising similar objections; see for example the
objections to the principle of “instrumental rationality” that are given by Joseph Raz (2011: chap. 8).

¹¹ So far as I know, I was the first writer to express this concern about whether the requirements of
rationality are any more than a mere “pretty mental pattern” (see Wedgwood 2003, 217).
As I explained above, I am assuming here that if “rationality” is a normative concept, then there is necessarily something good or worthy of commendation in any mental state or event that counts as rational. But we may well doubt whether it must always be true that there is something good simply in being coherent.

One way to raise these doubts is by focusing on the connection that is widely believed to hold between what is good and what there is a reason for pursuing. So, if there is always something good or worthy of commendation simply in being coherent, then it seems to follow that there is always a reason for us to be coherent. This idea that there is always a reason for us to be coherent is also more directly supported by the idea that rationality is a normative concept, and the requirements of rationality are fundamentally requirements of coherence, if it is also correct (as many philosophers claim) that normative concepts are those that are conceptually connected to reasons.12

It is controversial exactly what factors ground the various reasons that exist. On Joseph Raz’s (2011: 36-47) view, reasons for choice and action are grounded in values, while reasons for belief are necessarily “truth-related” in some way. Other philosophers, like Mark Schroeder (2007), view all reasons as grounded in desires: on this view, a fact is a reason for an agent to φ if and only if this fact is part of an explanation of why φ-ing will help to satisfy some of the agent’s desires.

The problem is that none of these familiar views of reasons makes it easy to see why there would always be a reason for us to be coherent. Whatever exactly the standards of coherence may be, coherence does not seem to have any clear connection to the factors that ground reasons for action or reasons for belief. Coherent sets of beliefs are not guaranteed to consist of beliefs in true propositions, or even to have a greater objective chance of involving beliefs in true propositions than some incoherent sets of beliefs; coherent plans and intentions are not guaranteed to do better than incoherent plans at leading to the agent’s promoting or respecting the values that matter in life; and coherent attitudes need not do better than some sets of incoherent attitudes at leading to the satisfaction of the agent’s desires.

For these reasons, then, it remains puzzling why there should be any genuine reason for us to be coherent. Since it is unclear whether there is always any reason for us to be coherent, it is equally unclear whether there is always anything good or worthy of commendation in

12 Compare Raz (2011: 85), who defends “the view that all normative phenomena are normative in as much as, and because, they provide reasons or are partly constituted by reasons.” As he says, this view “makes the concept of a reason key to an understanding of normativity.”
being coherent. As I have argued, however, it is unclear how rationality can require anything more of us than various forms of coherence. Given the assumptions that we are making here, this casts yet more doubt on the idea that rationality is a normative concept.

(e) Conclusion

We face a dilemma. On the one hand, there are compelling reasons for thinking that rationality is a normative concept; but on the other hand, the view that it is a normative concept faces all of the objections that we have just canvassed.

In the rest of this book, I shall defend the view that “rationality” is indeed a normative concept, against these objections. This defence of the view that “rationality” is a normative concept will reveal some of the concept’s most fundamental and important features, in a way that will take us a long way towards understanding what the norms of rational belief and rational choice really are.