13 Doxastic Rationality[1]

Ralph Wedgwood

It is widely accepted that the terms "justified" and "rational" can be used to express very closely related concepts. Unsurprisingly, then, it is not hard to find a way of expressing the distinction between "propositional" and "doxastic" justification in terms of "rationality":

(i) There is sufficient *propositional* justification for you to believe a proposition if and only if *it is rational* for you to believe the proposition.
(ii) You believe a proposition in a *doxastically* justified manner if and only if you *rationally believe* the proposition – that is, your belief in the proposition is *rationally held*.

An intuitive explanation of this distinction can be given as follows: The statement that *it is rational* for you to believe a proposition does not entail that you actually do believe the proposition. It entails only that the proposition is, given your cognitive situation, rationally suitable for you to believe. By contrast, the statement that you *rationally believe* the proposition does entail that you believe the proposition—and, moreover, that you believe it in a certain distinctively rational manner.

In most of the existing literature, this distinction between doxastic and propositional justification has been studied only in relation to *belief*. As it seems to me, however, this is far too narrow a view of the matter. The very same distinction applies to attitudes of many kinds. In this discussion, I am particularly interested in applying this distinction to partial *degrees of belief* or *credences*; but in principle the distinction also applies to many other attitudes—including practical attitudes, such as intentions and decisions. My goal here is to give an account of what it is for attitudes of all these kinds to have the property of being, as I shall put it, "rationally held."

Previous work on this distinction has been limited in another way as well. It has barely addressed the fact that, intuitively, both propositional and doxastic justification *comes in degrees*. Propositional justification clearly comes in degrees: among the beliefs that it is possible for you to hold now, there are some that it is *perfectly rational* for you to hold, some that it is *slightly* irrational for you to hold, and others that it is *grossly* irrational for you to hold. Similarly, doxastic justification seems also to come in degrees: some of the beliefs or other attitudes that thinkers hold are *more* rationally held than others. I shall also attempt here to show here how any account of these degrees of propositional justification can be used to give a corresponding account of degrees of doxastic justification.

In the first section below, I shall examine the most prominent account of doxastic justification, which appeals to the "basing" relation; as I shall argue, this account faces a series of grave objections. In the second section, I shall argue for an alternative account, which I shall call the "virtue manifestation" account—where the relevant virtue is precisely the virtue of *rationality*. This alternative account will be developed in more detail in the third section, where I shall give an account of how to measure the *degree* to which a particular attitude (such a belief or an intention) is rationally held. In the fourth section, I shall defend the account of the preceding section against some objections that some philosophers may be tempted to raise. Finally, in the fifth section, I shall highlight a further advantage of this account—a further puzzle to which this account can give an illuminating and satisfying solution.

The "Basing" Account

Many epistemologists claim that the crucial factor that differentiates a doxastically justified belief from beliefs that are not doxastically justified concerns the *basis* of the belief—what the belief is *based on*. Indeed, the very first philosopher to use the terms "propositional" and "doxastic" to mark this distinction was Roderick Firth (1978: 217), who claimed that for a belief to be "doxastically warranted," it has to be, as he put it, "psychologically based on or derived from the relevant evidence in a rational way."

A similar view is defended by evidentialists like Earl Conee and Richard Feldman (1985: 24)—although Conee and Feldman use the term "justified" solely for propositional justification, and employ the term "well-founded" to describe beliefs that Firth would call "doxastically warranted." In their view, belief in a proposition p is (propositionally) justified for a thinker at a time t if and only if having the attitude of belief toward p "fits" the "evidence" that the thinker has at t; and for the thinker's belief in p to be "well-founded"—that is, doxastically justified—not only must belief in p be justified, but thinker must also believe p "on the

basis of" evidence that the thinker has, where having this attitude "fits" that evidence.[2]

In my opinion, this account of doxastic justification is open to a series of grave objections. However, I shall not attempt to canvas all these objections here. Instead, I shall first argue that a belief's being "based on" evidence that the thinker has for that belief is neither necessary nor sufficient for doxastic justification. Then I shall argue that classical subjective Bayesianism is also extremely hard to reconcile with this "basing" account.

We might wonder at this point: What exactly is this "basing" relation?[3] Fortunately, I do not need to give a full answer to this question here. To some extent, we have a grasp of what the term "based on evidence" means in everyday English. So, to argue that basing is not sufficient, I need only to identify cases where it is intuitively highly plausible that a belief is based on evidence that supports it, but is not doxastically justified; and to argue that basing is not necessary, I need only to identify cases where it is intuitively highly plausible that a belief is not based on evidence, but is nonetheless doxastically justified.

First, then, the mere fact that a thinker holds a belief on the "basis" of evidence that supports that belief is not *sufficient* for the belief's being doxastically justified. This point is clearly shown by an example that is due to John Turri (2010: 317).[4] Consider a thinker whose evidence includes the two propositions p and "If p, then q." Suppose that from these two propositions the thinker infers q, and thereby comes to believe q. This seems to make it plausible that the thinker "bases" her belief in q on her beliefs in those two propositions, which together entail q. Since q is entailed by the thinker's evidence, it also seems that having the attitude of belief toward q "fits" or "is supported by" the thinker's evidence.

On reflection, however, it is clear that this is not sufficient for the thinker's believing q in a doxastically justified manner—because it does not guarantee that she infers q by means of *modus ponens*. Perhaps, as Turri suggests, she just has an insane disposition to infer *any* conclusion from any pair of premises whatsoever. While the particular inference that she accepts is an instance of modus ponens, it is also an instance of countless insane alternative rules as well. If she inferred q by following one of these insane rules, her belief in q would surely not be doxastically

---

2 Strictly, Conee and Feldman (1985: 24) allow that the evidence e on the basis of which the thinker believes p need not be the *total* evidence that the thinker has at the time, but e must be such that there is no more inclusive body of evidence e' that the thinker has at the time such that having the attitude of belief towards p does not "fit" e'. However, this complication will not matter for our purposes.

3 There has been much discussion of the basing relation. For an illuminating recent discussion, see Neta (2019).

4 For an earlier (though less memorable) presentation of this kind of argument, see Wedgwood (2002: 287).

justified—even though it is clearly based on evidence that supports it. Thus, being held on the basis of supporting evidence is not sufficient to make a belief count as doxastically justified.

Admittedly, this is only a counterexample to the formulation that is given by Conee and Feldman. It is not a counterexample to the formulation that is given by Firth, who says that for a belief to doxastically warranted, the belief must be "psychologically based on or derived from the evidence in a rational way"—and a thinker who is following an insane rule of inference is presumably not basing their belief on the evidence in a "rational way". For this reason, some philosophers—such as Neta (2019)—set out to inquire what more is required of an instance of the basing relation if this instance is to be a case of doxastic justification. As I shall argue in the next section, however, the most plausible account of what makes an instance of the basing relation a case of doxastic justification must appeal to a further feature—which turns out *also* to be present in cases in which a belief is not in any obvious sense "based" on evidence at all. On further reflection, it becomes plausible that it is this further feature—and not anything involving the "basing" relation—that is really essential to doxastic justification.

Even if this worry about the *sufficiency* of "basing" can be addressed, it is also doubtful whether being "based" on evidence is *necessary* for doxastic justification. One way to see this point is by coming to appreciate that there are some beliefs that seem not to be based on evidence at all, even though some of these beliefs seem clearly to be doxastically justified. In particular, there are three salient examples of beliefs that seem, at least *prima facie*, not to be based on evidence.

First, there are one's deeply entrenched background beliefs—beliefs that one has held for years, and which have by now just become an entrenched part of one's outlook on the world.[5] For example, I now have a belief about the name that my paternal grandmother was given at birth—specifically, I believe that her name was "Diana Hawkshaw." On what evidence is this belief now "based"? Clearly, I cannot now "base" this belief on any evidence that I do not now *possess*. Presumably, at some time while I was a child, I had an experience as of being told by some family member that my grandmother's name was "Diana Hawkshaw." However, since I now have absolutely no recollection of that experience, it seems that neither that past experience itself nor any recollection of that past experience is part of the evidence that I now possess.

Admittedly, I do now have various beliefs *about* this belief. For example, I now believe that my grandmother's name must have cropped up in conversations with family members on a number of occasions in the

5 The significance of these beliefs that are no longer based on evidence or reasons is discussed by Harman (1986, Chap. 4), Peacocke (1986, Chap. 10), and Millar (1991, Chap. 6).

past, and on all those occasions my belief that her name was "Diana Hawkshaw" was confirmed. But it is doubtful whether my belief about my grandmother's name is now "based on" these beliefs that I now have about that belief. It is not particularly plausible to say that the best explanation of why I now hold this belief is because I now believe that I must have had conversations in the past in which this belief was confirmed. It seems more plausible to say that I hold this belief now simply because it has become part of my entrenched system of background beliefs.

Some philosophers—such as J. L. Pollock and Joseph Cruz (1999: 48)—claim that this belief is based on my *seeming to remember that* my grandmother's name was "Diana Hawkshaw." But it is not clear that this state of *seeming to remember that* … is anything other than the belief itself, stored away and then accessed from long-term memory. So, it seems dubious to look for a "basis" for the belief in such a state. For these reasons, then, it seems more plausible to conclude that this belief is not now "based" on any evidence at all. The same is true of countless entrenched background beliefs about history, geography, language, science, and much more. But it seems plausible that I hold almost all these beliefs in a rational and doxastically justified manner.

A second example of beliefs that cannot be based on evidence would only arise on certain conceptions of "evidence"—namely, conceptions on which there are some propositions that form part of one's evidence precisely because one believes or knows them. This is admittedly not true on all conceptions of evidence, but it is true on some well-known conceptions, such as that of Timothy Williamson (2000, Chap. 9). So, suppose that a proposition p is part of my evidence, and it is part of my evidence in virtue of the fact that I believe p. As I have noted, we are not assuming any complete account of the "basing" relation here; but it seems very odd to suggest that I "base" my belief in p on my belief in p—or even that I base it on a large collection of beliefs that includes my belief in p. In general, then, it seems that no proposition that is part of one's evidence in virtue of one's believing it can itself be believed on the basis of one's evidence.

A third example concerns *logical truths*. On some conceptions of what it is for evidence to "support" believing a proposition, every proposition that is *entailed* by one's evidence is supported by one's evidence. But of course, logical truths are entailed by every body of evidence whatsoever. So, on these conceptions, logical truths are trivially supported by all evidence whatsoever. For this reason, it seems doubtful whether in believing such a logical truth, I believe it because of the *specific* evidence that I have, given that I would have just as much justification for believing this logical truth whatever evidence I had had.

While the proponent of the "basing" account might be able to resist counterexamples of one of these three kinds, the fact that counterexamples arise in so many different areas poses a severe challenge to this account.

In addition, it also seems that classical subjective Bayesianism about rational belief is hard to reconcile with the "basing" account.[6] On this Bayesian view, for every pair of times t_0 and t_1, and for every perfectly rational thinker x, if the conjunction of all the evidence that x has acquired between t_0 and t_1 is e, then at t_0 and t_1 x holds probabilistically coherent systems of credences C_0 and C_1 respectively, such that for every proposition p for which C_0 and C_1 are defined, $C_1(p) = C_0(p|e)$. That is, the thinker's current credences are the result of *conditionalizing* her earlier credences on the conjunction of all the evidence that she has acquired since then. (Strictly speaking, this Bayesian view only applies to thinkers who do not *lose* any evidence over the relevant period of time.)

It is clear on reflection that this classical subjective Bayesian view should be taken as a view of abstract "propositional" rationality. Nothing in the view prevents it from being possible that a thinker might conform to the requirements of this Bayesian theory through a remarkable cosmic accident.[7] First, the thinker might through some lucky accident have a probabilistically coherent system of credences at t_0, and then through a second still more extraordinary accident shift at t_1 to the result of conditionalizing that first system of credences on all the evidence that she has acquired between t_0 and t_1. If the thinker's conformity to these Bayesian requirements is a sheer lucky fluke in this way, then the credences that she has at t_1 may be the credences that it is abstractly or propositionally rational for the agent to have, but they are not rationally held—or, in other words, they are not doxastically justified.

In general, if the "basing account" is the right conception for full or outright beliefs, it is surely also the right conception for credences or partial degrees of belief: it seems much more plausible that a unified account of doxastic rationality can be given—an account that applies both to full beliefs and to partial degrees of belief—than that radically different accounts apply to beliefs of these two different kinds. So, we should consider a "basing account" of the doxastic rationality of systems of credences as well. According to this account, for a system of credences to be rationally held by a thinker at a time is for this system of credences to be "based" on the evidence that the thinker has at the time.[8] However, there is an obvious problem with this account.

It is a crucial feature of this Bayesian view that, for every time, there are *two* factors that explain why it is rational for the thinker to have a particular system of credences at that time: (a) the prior credences that

6 For a classic statement of this kind of subjective Bayesianism, see Jeffrey (2004).
7 For this point, see Staffel (2019: 129f.). Relatively few epistemologists have discussed in any detail what it is for credences to be doxastically justified; for an interesting exception, see Dogramaci (2018) and Tang (2016).
8 For an account of doxastically justified credences that is broadly along these lines, see Smithies (2015).

she had in the past, and (b) the evidence that she has acquired since then. The "basing account" mentions that second factor (b), the evidence that the thinker has at a time, but signally omits the other factor (a) that Bayesianism appeals to—the prior credences that the thinker had in the past.

Thus, for these credences to be rationally held, it is not enough that (i) the thinker bases her credences on the evidence that she has, and (ii) it is rational for her to have these credences at that time. It could be that the thinker bases her credences on her evidence, without manifesting any disposition to conform to the principle of conditionalization, and it is simply an extraordinary cosmic fluke that on this occasion the system of credences that she happens to base on the evidence is a system that results from conditionalizing her past credences on the evidence.

The proponents of the "basing" account might try to suggest at this point that what they mean by "evidence" is something *broader* than what the Bayesians normally mean. In particular, they might try to suggest that, in their terminology, the Bayesian view implies precisely that the thinker's past credences constitute part of her current evidence. But this distorts the ordinary meaning of the term "evidence" beyond its breaking point. Specifically, this suggestion implies that on the Bayesian view, a rational thinker always "bases" her present attitude toward p at least in part on the attitude that she had toward p in the past—and, moreover, that she often does so when she has no current memory of what that past attitude was. This is surely not an instance of what we ordinarily mean by phrases of the form "based on one's current evidence for p."

The proponent of the "basing account" might make one final move at this point. They might amend their account, in the following way. According to this amended account, the system of credences that you now have is rationally held if and only if it is based on *both* the evidence that you now have *and* your past credences. However, it is doubtful whether it can really be the same "basing" relation that holds both (a) between your current credences and your current evidence and (b) between your current credences and your past credences. Even if it is the same relation, it is not clear why this "basing relation" cannot hold as a result of an extraordinary cosmic accident. So, it is doubtful whether this attempt to rescue the "basing" account can succeed in avoiding these objections.[9]

For all these reasons, then, it seems to me advisable to abandon the "basing" account, and to look for an alternative approach. Fortunately,

---

9 A second potential problem would arise if Worsnip (2018) is correct, and the credences that fit optimally with the agent's "evidence" might not be ideally coherent. On this view, the credences that are "based" on the agent's evidence might not even be probabilistically coherent—which would clearly prevent them from being perfectly rational by Bayesian standards.

an alternative approach is close to hand—in an idea that I have defended elsewhere, that rationality is a *virtue*.

Rationality as a Virtue

What is the core of the distinction between propositional and doxastic justification—that is, between the beliefs that *it is rational* for me to hold and the beliefs that I *rationally* hold?

Part of the distinction, evidently, is this: if it is rational for me to hold a certain belief, it does not follow that I actually hold the belief—whereas if I rationally hold a certain belief, it trivially follows that I hold the belief. However, even if a thinker *believes* a proposition *p* at the same time as *p*'s being a proposition that *it is rational* for her to believe, this is still not sufficient for her to count as *rationally believing p*. It might simply be a lucky fluke that both these two conditions hold at the same time. In general, then, for the thinker to count as rationally believing the proposition, it must *not* be lucky fluke that these two conditions hold. It must somehow be *no accident* that the thinker believes the proposition at the same time as its being a proposition that it is rational for her to believe.

As I have argued elsewhere (Wedgwood 2017: 140–2), this distinction is in fact precisely analogous to a distinction that Aristotle drew in the *Nicomachean Ethics* (1105a17–b9). An agent might perform an act of type *A* at the same time as being in a situation in which *it is just* for her to perform an act of type *A*—even if it is simply a lucky fluke that both of these conditions hold (perhaps it just so happens that her wickedly unjust plans require her to perform an act of type *A* in this situation). In this case, the agent might be doing a *just act*, but she would not count as *acting justly*.

If an act is to be a case of the agent's acting justly, the following conditions must hold. First, in performing this act, the agent must be manifesting an *appropriate disposition*—specifically, a disposition that non-accidentally tends to result in the agent's doing just acts. Second, this disposition must manifest itself in its characteristic way—so that the act in question really is just (at least to a significant degree). Some theories of virtue would add further conditions (for example, perhaps the manifestation of this disposition must take the form of the rational pursuit of goals—such that there is a reliable tendency for the rational pursuit of these goals to result in the agent's doing just acts). But we need not worry about these details here. The important point for our purposes is that having such an appropriate disposition is part of what is involved in possessing the *virtue of justice*—that is, part of what is involved in being a just person.

In a similar way, I propose, if a belief is a case of the thinker's rationally believing *p*, the following conditions must hold. First, in believing *p*, the thinker must be manifesting an *appropriate disposition*—specifically, a

disposition that non-accidentally tends to result in the thinker's believing propositions that it is rational for her to believe.[10] Second, the disposition must manifest itself in its characteristic way—so that the belief in question really is one that it is (to a least a significant degree) rational for the thinker to hold. Such a disposition is an example of a broader genus—namely, dispositions that non-accidentally tend to result in the thinker's having attitudes that it is rational for her to have. Having such dispositions seems to be part of what is involved in being a rational thinker. I shall refer to dispositions of this kind as *rational dispositions*. My central proposal here, then, is that for an attitude to be rationally held is for it to be a characteristic manifestation of a rational disposition.[11]

In this way, there seem to be three features associated with rationality:

(i)   Being an attitude that *it is rational* for the thinker to have at the time—a feature exemplified by some of the possible attitudes that are available for the thinker to have at the time.

(ii) *Having a rational disposition*—a feature exemplified by thinkers, which non-accidentally tends to result in the thinkers' having attitudes that have the first feature.

(iii) An attitude's being a case of *the thinker's thinking rationally*—a feature exemplified by an attitude whenever the attitude is the manifestation of a disposition that makes it the case that the thinker has the second feature, and the disposition is manifested in its characteristic way, so that the attitude has the first feature (at least to a significant degree).

As we have seen, there is a parallel trio of features associated with justice:

(i)   Being an act that *it is just* for the agent to perform at the time—a feature exemplified by some of the possible acts that are available to the agent at the time.

(ii) Having a *disposition* that non-accidentally tends to result in one's performing such just acts—a feature exemplified by agents.

---

10 What if a belief results from the operation of *two* different dispositions—one of which is rational, and the other irrational? The question here is whether (a) both these dispositions are causally necessary parts of a single sufficient causal explanation, or (b) each of these dispositions is sufficient by itself (so that we have a case of causal overdetermination). I am inclined to think that in the former case, the belief is *not* a case of believing rationally, but in the latter case, it *is* a case of believing rationally. So, more precisely, what is required for believing rationally is that the agent must manifest rational dispositions, and the manifestation of these rational dispositions must be a sufficient explanation of the agent's having the belief. But it is not required that this should be the *only* sufficient explanation of the belief.

11 For a related view, developed within a distinctively externalist framework, see Lasonen-Aarnio (forthcoming).

(iii) An act's being a case of the agent's *acting justly*—a feature exemplified by an act whenever the act is the manifestation of a disposition that makes it the case that the agent has the second feature, and the disposition is manifested in its characteristic way, so that the act has the first feature (to a significant degree).

Each of these three features associated with justice is an evaluatively positive feature—a feature in virtue of which the act or agent in question is good in a certain respect. In the case of rationality too, each of these three features seems to be an evaluatively positive feature—a feature that makes the attitude or thinker in question good in certain respect. Because of these parallels between rationality and a paradigmatic virtue like justice, it seems reasonable to conclude that rationality itself is a virtue. These rational dispositions are at least part of what is involved in possessing the virtue of rationality—or in other words in being a rational thinker.

I have suggested that having rational dispositions of this kind is "part" of what is involved in being a rational thinker. That is, having such dispositions is a constitutively necessary condition, but perhaps not a sufficient condition, for being a rational thinker. This is for two reasons. First, to count as a "rational thinker", it is not enough just to possess a disposition of this kind that covers an extremely narrow range of cases; one has to possess such dispositions covering a sufficiently wide range of cases. Second, in most contexts, it would not be true to describe someone who possessed such rational dispositions as a "rational thinker", unless these dispositions were actually *manifested*, at least much of the time, in the thinker's mental life. As we shall see in Section 13.4 below, it is possible to possess a disposition even if the disposition is blocked or inhibited from being manifested in one's actual thinking by the counteracting influence of some interfering factor.[12]

Exactly how these dispositions operate depends on contingent facts about how the human mind works, or about the particular cognitive skills and abilities that the thinker has acquired. For example, in many cases they may involve what cognitive scientists think of as mental "heuristics", of the sort that are characteristic of quick-and-dirty "System 1" thinking.[13] I am inclined to suspect that for our purposes it does not matter exactly how these dispositions work, so long as these dispositions non-accidentally tend to result in the thinker's having attitudes that it is rational for her to have. Whenever one of your beliefs is the manifestation of rational dispositions of this kind, then in a clear sense it is no accident that it is a belief that it is rational for you to have. In this case,

---

12 For this point about dispositions, see Bird (1998).
13 For a discussion of the distinction between "System 1" and "System 2" thinking, see Kahneman (2011).

this is not just a belief that it is rational for you to have—it is a belief that is actually rationally held.

Moreover, so far as I can see, this conception of what it is for a belief to be rationally held need not presuppose any particular conception of belief. All that is presupposed is that your believing proposition p at time t can itself count as your manifesting a certain disposition that you now have. This presupposition is compatible with many different conceptions of belief. Even if your believing p at t consists in the fact that a sentence of your "Language of Thought" that means p is tokened in your "Belief Box" at t, this fact might itself count as your manifesting an appropriate disposition—specifically, a disposition for sentences of the relevant kind to be tokened in your "Belief Box" in response to stimuli of a certain corresponding kind.

To fix ideas, in what follows I shall assume a more detailed conception of dispositions.[14] Specifically, I shall assume that every disposition can be specified by a *function* that maps *stimulus* conditions onto *response* conditions. For example, the disposition of *fragility* can be specified by a function that maps the stimulus condition *being struck at t* onto the response condition *shattering shortly after t*. The notion of a disposition also presupposes the idea of a *range of normal cases*. These are actual and merely possible cases in which other things are equal—cases in which factors that would interfere with the operation of the disposition are absent. Something possesses the disposition if and only if it has intrinsic properties in virtue of which, in any *normal* case in which it is in one of these stimulus conditions, it goes into the corresponding response condition. For example, something is fragile if and only if it has intrinsic properties in virtue of which, in any normal case in which it is struck, it shatters shortly afterwards.

Each of the rational dispositions that we are interested in is a mental disposition of a certain kind. Specifically, the stimulus conditions for each of these dispositions concern the facts about one's cognitive situation that are beyond the control of one's reasoning capacities at the relevant time—such as facts about one's past attitudes, and facts about the sensory experiences, memories, and emotions that one has at the time. The disposition's response to these facts consists in one's having a certain set of attitudes at the time. The disposition can be specified by a function that maps stimulus conditions of this kind onto corresponding response conditions.

The notion of "manifesting" a disposition is a broadly causal notion. A disposition is triggered by one of the relevant stimulus conditions; and when the disposition is manifested, the response is caused by the stimulus. For example, one rational disposition that a thinker might have

---

14 I am drawing here on some of my earlier discussions of dispositions; see Wedgwood (2007: 27f., and 2017: 76–8).

is what we could call the *modus ponens* disposition. This disposition is triggered by the thinker's considering an argument that is an instance of *modus ponens*, and is manifested in the thinker's accepting that argument (at least in the sense of conditionally believing the argument's conclusion, conditionally on the assumption of the argument's premises). In manifesting this disposition, the response (the thinker's acceptance of the argument) is caused by the stimulus (the thinker's considering the argument).

If this account is along the right lines, there is a *causal* element in an attitude's being doxastically justified or rationally held: whether or not one believes p in a doxastically justified manner depends on the causal explanation of one's believing p—on the dispositions that one manifests in holding this belief.[15] By contrast, there is no such causal condition on propositional justification. Whether or not one has propositional justification for p just depends on how the attitude of believing p "fits" with one's cognitive situation. Even if one does actually believe p, the fact that one has propositional justification for believing p entails nothing about the causal explanation of why one believes p.

It is because of this causally explanatory element that we can also talk about such doxastically justified rational thinking as consisting in having attitudes "for the right kind of reasons." The "reasons" that we are alluding to here are the *motivating reasons* for which one has the attitude in question—that is, the intuitively intelligible psychological explanation of why one has the attitude. (We may assume that these psychological explanations always appeal at least implicitly to the dispositions that the thinker is manifesting in having the attitudes in question.) If the attitude is rationally held, then this psychological explanation is "of the right kind" in the sense that it is the right kind of explanation to make the attitude into a manifestation of rational dispositions of the kind that I have described.

Degrees of Rational Virtue

The account sketched in the previous section is, in an important respect, rough and incomplete. It does not take account of the fact that rationality comes in *degrees*. Indeed, all three of the good features that I distinguished above—the abstract "propositional" rationality of attitudes, the rationality of the relevant dispositions, and the property of being an attitude that is rationally held—seem to come in degrees. Of the various attitudes toward p that are now available to you, some are more rational than others; some mental dispositions of the relevant kind are more rational than others; and some attitudes are more rationally held than others.

15 Keith Lehrer (1974: 125) argued against imposing any causal requirement on doxastic justification. For a largely convincing reply to Lehrer, see Audi (1983).

On reflection, it is clearly important to give an account of these degrees of rationality. This is because it is extremely unlikely—if not downright impossible—for creatures like us to have dispositions that are *ideally* rational, in the sense that these dispositions *infallibly* result in *perfectly rational* attitudes in *all* normal cases in which they are manifested. If this sort of ideal rationality is effectively unattainable for us, what we need is some way of comparing which of the actually available dispositions are more rational and which are less rational.

Let us start by investigating the degree of rationality that a mental disposition may have *relative to a range of cases*. This range of cases may consist of *all* normal cases in which the disposition is manifested at all, or it may consist only of some narrower range of cases, all of which are sufficiently similar to some particular case that is under consideration. (We shall later consider how to make sense of statements that character-ize dispositions as having some degree of rationality that is not explicitly relativized to a range of cases in this way.)

What is it, then, for one disposition to be more rational than another (relative to to a given range of cases)? The rough idea is that the more rational disposition is more *reliable* at yielding attitudes that it is abstractly or propositionally rational for the thinker to have. This notion of a "more reliable" disposition may be made more precise in the following way.

Crucially, as I noted at the outset, the abstract propositional rationality of attitudes also comes in degrees. Some of the available attitudes "fit" the agent's cognitive situation better than others. For example, given my past beliefs and my current sensory experiences, the attitude of *believing* that I am now sitting in a chair "fits" my cognitive situation much better than the attitude of *doubting* or *disbelieving* that I am now sitting in a chair.

Moreover, I shall assume that the degree to which a set of attitudes is abstractly or propositionally rational can in principle be *measured*. This assumption is especially plausible in the case of credences or partial degrees of belief. According to many probabilistic theories in formal epis-temology, for every thinker and every time there is some probability dis-tribution such that the degree to which a set of credences is rational for that thinker and that time is determined by how *closely* those credences *approximate* to that probability distribution.[16] To fix ideas, we may assume that some probabilistic theory of this sort is correct—although strictly all that we need for our purposes is the more general assumption that the degree to which a set of attitudes is rational can in principle be measured.

As I explained, we are investigating the degrees of rationality that dis-positions can have relative to a certain "range of cases." This "range of

---

16 For a pioneering discussion of how to measure the propositional rationality of cre-dence assignments, see Staffel (2019).

cases" can be thought of as a *set of possible worlds*—which are all alike in that the thinker manifests the disposition at the relevant time, but differs from each other in the precise cognitive situation of the thinker at the time, and in the response that the manifestation of the disposition yields to that situation. At each of these worlds, the thinker's response has a certain degree of abstract propositional rationality. The facts of the thinker's psychology also fix a certain *chance* function, which assigns a conditional chance to each of these worlds, conditionally on the relevant disposition's being manifested at the relevant time.[17]

Together, these degrees of abstract propositional rationality and this chance function determine the *expected degree of rationality* of the responses that the disposition yields within this range of cases.[18] I propose that for one disposition to be *more rational* than another (relative to a range of cases) is for the first disposition to have a higher expected degree of rationality (within this range of cases).

According to this proposal, then, for a disposition to be *ideally* rational within a given range of cases is for it to be conditionally *certain* that the thinker's attitudes will be *perfectly rational* in all these cases, given that the disposition is manifested at the relevant time. Among the many dispositions that are less than ideally rational, some are more rational than others. For example, it might be conditionally certain that the attitudes resulting from one disposition D_1 will be 80% rational, given that this disposition D_1 is manifested, while there might be a conditional chance of 50% that the attitudes resulting from a second disposition D_2 will be 90% rational and a conditional chance of 50% that these attitudes will be 60% rational. In this case, the expected rationality of the attitudes that result from D_1 is 80% while the expected rationality of the attitudes that result from D_2 is 75%. According to my proposal, then, D_1 is more rational than D_2.

For the rest of this discussion, however, I need not assume the precise details of this account of what it is for one disposition to be more rational than another, relative to a range of cases. These details are given here only as a proof of concept, to make it plausible that there are no insuperable difficulties in developing such an account; I need not deny that other accounts may be worth considering.

---

17 If, as Glynn (2010) has argued, chance functions are relativized to levels of causal structure, these chance functions must be relativized to the *psychological* level. For discussion, see Wedgwood (2017: 80f.).

18 Strictly, this requires that the degrees of propositional rationality of the disposition's outputs in the relevant cases can all be measured on the same scale. On one view of the matter, these degrees of rationality are only comparable between credence assignments that are defined over the *same* set of propositions. If this view is correct, the relevant "range of cases" where the rationality of the disposition is assessed must be restricted to cases where the thinker considers the same set of propositions.

How can this account of the degrees of rationality that dispositions have *relative to a range of cases* make sense of statements that characterize dispositions' degrees of rationality *without* any explicit relativization to ranges of cases? Somehow, the meaning that such a statement has, along with the context in which the statement is made, must determine some particular range of cases as the range that is *relevant* to the statement's truth-conditions in that context. In fact, I do not need to commit myself here to any particular account of how the statement's meaning together with the context determine this. Again, however, it may be useful to gesture in the direction of an account—to make it plausible that no special obstacles stand in the way of developing such an account.

Two possible accounts stand out as particularly promising. According to the first account, if a statement is not explicitly relativized to a range of cases, it is always the *widest* range of cases—including all normal possible cases in which the disposition is manifested at all—that is relevant to the statement's truth-conditions. By contrast, the second account is a form of *contextualism*. On this contextualist account, it is part of the *context* in which the rationality of a disposition is discussed that the participants in the conversation have a certain range of cases at least roughly in mind. In some contexts, they may be focusing on the widest range—all normal possible cases in which the disposition is manifested, but in other contexts, they may be focusing on some *narrower* range of cases that is somehow salient. According to this second account, it is this contextually salient range of cases that is relevant to the truth-conditions of such statements.[19]

At all events, it is straightforward to extend this account of the degrees of rationality that *dispositions* may have, to provide an account of the degrees of rationality of *manifestations* of those dispositions. In saying that one attitude is "more rationally held" than another, we are in effect saying that the first attitude is the manifestation of a more rational disposition than the second.

Finally, we can also use this account of degrees of rationality to give an account of what it means to make the non-comparative statement that a belief is "rationally held" (or that it is "well founded" or "doxastically justified"). I propose that this statement means simply that the belief is the manifestation of a *sufficiently rational* disposition.

Many philosophers of language and semanticists would also accept a contextualist interpretation of this talk of what is "sufficiently rational." On this contextualist interpretation, we can always ask, "The disposition's degree of rationality is sufficient *for what*?"—and an answer to this question must somehow be implicit in the context, if the statement that

---

19 In general, I would argue that disposition ascriptions are typically context-sensitive in this way. For a related claim about the context-sensitivity of explanatory claims, see Wedgwood (2020: Section 5).

the belief is "rationally held" is to have determinate truth-conditions.[20] Again, however, I need not commit myself to this second kind of contextualism here. Perhaps the normal meaning of "rationally held" fixes a certain degree of rationality as "sufficient" independently of the conversational context: for example, perhaps we can say, in general, that a belief counts as "rationally held" if and only if it is the manifestation of a disposition that exemplifies as high a degree of rationality as is feasibly achievable for the majority of human beings. At all events, it seems that there is no difficulty in principle in using this account of degrees of rationality to give an account of what makes it true to say that an attitude such as a belief is rationally held.

Objections and Replies

The account sketched above may seem reminiscent of *reliabilist* approaches to epistemology, such as the approach of Goldman (1979). So, it is natural to inquire whether it is open to the same objections as reliabilism.

First, however, before inquiring into this, we need to see that the account differs from the familiar forms of reliabilism in a crucial way. The most familiar forms of reliabilism focus on reliability at (a) generating *true* beliefs and (b) not generating *false* beliefs. In other words, the relevant kind of "reliability" for these familiar forms of reliabilism is *truth-conduciveness*. This kind of reliability or truth-conduciveness lies at the center both of the older kind of reliabilism that was developed by Goldman (1979), with his emphasis on the idea of "reliable processes," and of the more recent "virtue epistemology" of Sosa (2007).

By contrast, my kind of reliabilism about the virtue of rationality focuses on a different kind of reliability—not truth-conduciveness, but *rationality-conduciveness*. On my account, for an attitude that the thinker holds to be rationally held, the attitude must be the manifestation of an appropriately "reliable" disposition; but the disposition needs to be "reliable," not at resulting in *true* beliefs, but at resulting in the thinker's having credences that it is *rational* for the thinker to have.[21]

For this reason, my account is quite compatible with epistemological *internalism*. This account is thus compatible with the view that the degree

20 For a contrary view about "rational," see Siscoe (2021). According to Siscoe, "rational" always strictly means "perfectly rational" (just as "certain" strictly means "fully certain" and "clean" means "completely clean")—although this strict meaning can sometimes be used loosely or imprecisely.

21 It is unclear how to the notion of reliability as truth-conduciveness can even be applied to partial credences or degrees of belief. For an illuminating exploration of this issue, see Tang (2016), although from my internalist point of view, Tang's way of assessing credences' reliability mixes together notions—like truth and evidence—that are best kept more cleanly separated.

to which each of the various available credence assignments is rational for a thinker to have at a time is determined by purely "internal" facts about the thinker's mind at that time. In this way, my account is compatible with internalism about abstract "propositional" rationality.

Indeed, my account is even compatible with a kind of internalism about rationally held (or doxastically justified) attitudes. It is admittedly not plausible that we can infallibly introspect either (a) which dispositions each of our beliefs or other attitudes result from, or (b) how reliable these dispositions are at resulting in our having attitudes that it are rational for us to have. But if internalism about abstract propositional rationality is true, then both of these facts (a) and (b) are, broadly speaking, facts about our minds that are independent of any facts about our environment that could vary while our minds continue to function in the same way.

A further difference from some forms of reliabilism is that my account does not appeal to the notion of a "reliable *process*." The process as a result of which the thinker holds a belief reaches back into the past – sometimes very far back. The account that I have proposed does not focus on this historical process, but on the explanation of why the thinker holds the belief in question at the particular time in question—where I assume that this explanation will be equivalent to one that explicitly cites the dispositions that the thinker manifests in holding this belief at this time.

Nonetheless, some philosophers might object to the appeal to any kind of "reliability." Specifically, these philosophers might object that the "reliability" of a disposition is a matter of its *track record*—whereas it seems that the fact that a particular belief is rationally held at a particular time does not depend on the track record of any of the thinker's traits, but purely on how things are with that belief at that time.

However, this objection misinterprets the kind of "reliability" that my account appeals to. This kind of reliability does not depend on the track record that of any of the thinker's traits have *over time*. It depends on how the disposition performs across a range of *possible worlds*. Talking about the reliability across these possible worlds of the disposition that one manifests in holding a given attitude is really just a way of talking about the *modal* properties that this disposition has in this case; specifically, it is a just of discussing whether in this case the disposition *safely* or *robustly* or *non-accidentally* results in one's holding an attitude that it is rational for one to hold. This is a feature of how the belief is held at the particular time in question—not a feature of any kind of "track record."

Similarly, even the ascription of this mental disposition to a thinker is not a claim about any kind of "track record." There are plenty of dispositions that an item may have without ever manifesting these dispositions. One reason for this is that the item may never be in the relevant stimulus condition: certain poisons may never be ingested; certain fragile objects

may never be struck. Moreover, even if the item is sometimes in the relevant stimulus condition, it may never be in such a stimulus condition in a *normal* case: for example, it may be that a certain poison is only ever ingested by someone who has taken the antidote. In other words, it may be that interfering factors always intervene to block the manifestation of the disposition. According to the assumptions about dispositions that I articulated in Section 13.2 above, to ascribe a disposition to a thinker is not to make any claim about the thinker's track record in the actual world, but to make a claim about the modal properties of some of the thinker's traits with respect to normal cases—cases where interfering factors are absent. For these reasons, then, this objection to my account of rationally held attitudes is misplaced.

The best-known objection to reliabilism is the so-called "generality problem," which was pressed by Conee and Feldman (1998). Consider the version of reliabilism that says that a belief is rationally held just in case it results from a belief-forming process that is of a *generally reliable type*. The trouble with this is that every token process that results in a belief is an instance of countlessly many types—and some of these types are much more reliable than others. How are we to tell which, of the countlessly many types that a given token process instantiates, is the type whose degree of reliability determines whether or not the belief is rationally held? The theory as stated seems to have no principled way of answering this question.

However, as I shall now explain, there are several differences between my account and the version of reliabilism that Conee and Feldman criticize, which enable my account to answer this problem.

First, unlike the notion of a "belief-forming process," the very notion of a "disposition" introduces a kind of generality. As explained in Section 13.2 above, the disposition is specified by a *function* from stimulus conditions to response conditions—where both the stimulus conditions and the response conditions are in effect general *types* of condition (not particular conditions). Thus, the very nature of the disposition determines a certain range of cases—namely, the range of normal cases in which the disposition is manifested. These cases are all similar to each other along two dimensions: (a) in the dispositions that are manifested in these cases—and so in the stimulus and response conditions that these dispositions involve, and also (b) in the factors that make these cases count as normal.

Second, as we saw in the previous section, my account makes room for a kind of *contextualism* about what it means to ascribe a degree of rationality to a disposition. It may not always mean that the disposition results in a rational attitude in *every* normal possible case in which it is manifested. It may be that, in the context in which a degree of rationality is ascribed to the disposition, a narrower range of cases is salient. For example, in such contexts, our focus may be on a particular belief that

a thinker holds, and on the range of cases that are sufficiently similar to *this* belief on the two relevant dimensions of similarity—namely, (a) the manifestation of the relevant dispositions and (b) the factors in virtue of which these cases count as normal. For it to be true to ascribe a certain degree of rationality-conduciveness to the disposition in a given context, the disposition must yield beliefs with corresponding degrees of rationality in the range of cases that is relevant in this context.

In general, these two elements—the inherent generality of the dispositions that are manifested (along with the two dimensions of similarity associated with these dispositions), and the kind of contextualism that I have described—seem to provide enough material for a successful response to the generality problem.

A Further Advantage of this Account

Probabilism is often thought to be inadequate as an epistemology for mathematics. Presumably, every provable mathematical truth has probability 1. So, the only perfectly rational level of confidence that you can have in a mathematical truth is the highest possible level—that is, credence 1, the attitude of being totally convinced of the mathematical truth. However, consider an extremely complex proposition p that is in fact a mathematical truth. Professional mathematicians may often have a level of confidence in this proposition p that is considerably lower than credence 1. Can it really be right to say that it is not rational for these mathematicians to have such non-maximal levels of confidence in this truth? Considerations like these have often led philosophers—especially philosophers of mathematics—to doubt the correctness of probabilistic approaches to epistemology.[22]

The account given above provides a way of defending probabilism against this objection. Credence 1 is indeed the only attitude that one can take toward this mathematical truth that it is *perfectly* rational for one to have. But it may also be that a lower level of confidence in this truth is the only attitude that would result from any of the mental dispositions that are both (i) sufficiently rational and (ii) genuinely available to the normal human thinker.

So, we can interpret the judgment that it is rational for the mathematician to have a non-maximal degree of belief in the mathematical truth, not as a judgment about abstract "propositional" rationality, but as a judgment about the attitudes that would result from any sufficiently rational available disposition. To interpret the judgment in this second way, we evidently need some way of making sense of what it is for dispositions to

---

22 I have encountered this argument in conversation with both Christopher Peacocke and Ian Rumfitt.

be "sufficiently rational." Fortunately, we can make sense of this by using the account that I sketched in Section 13.3.

Why would it be that the most rational available dispositions would not result in the mathematician's instantly assigning credence 1 to this complex mathematical truth? The reason is just that, if the mathematician were instantly to assign credence 1 to this truth, as soon as she considers it, this could only be through sheer dumb luck; it could not be through the operation of a reliable disposition.

To explore some of the implications of this, let us consider an example that is due to Dogramaci (2018). Before you get a chance to do the math, a credence of 0.1 that the trillionth digit of π is a 2 seems entirely rational. But of course, the probability that the trillionth digit of π is a 2 must be either 0 or 1. So, according to a probabilistic theory of abstract propositional rationality, a credence of 0.1 in this proposition is necessarily irrational: the only perfectly rational attitude to have toward this proposition is either credence 0 (if it is false) or credence 1 (if it is true).

However, consider the disposition that one would be manifesting in having a credence of 0.1 in this proposition. Presumably, this disposition would also yield a credence of 0.1 in each of the nine other propositions of the same form—that the trillionth digit of π is a 3, that it is a 4, and so on. To keep things simple, let us assume that this disposition is *certain* to yield this credence assignment in every relevant case. So, this disposition yields an assignment of credence 0.1 to all ten propositions—where, of these ten propositions, nine have probability 0 and one has probability 1. On plausible measures of degrees of rationality, this credence assignment is at least *more* rational than (say) an assignment of credence 0.5 to all of these ten propositions.[23] Thus, this disposition is more rational than an alternative disposition that is certain to yield an assignment of credence 0.5 to each of these propositions in every relevant case.

In general, it seems that every disposition that is available to us will either assign the *same* credence to all of these ten propositions, or else it will just assign credences to them at random. Thus, the most rational available disposition seems to be the one that assigns a credence of 0.1 to all of these propositions—in spite of the probabilistic incoherence of these credence assignments.

In most discussions of the distinction between propositional and doxastic justification, it is assumed that doxastic justification entails propositional justification. However, the example just discussed shows that this is not true in general. If as a matter of fact the trillionth digit of π is a 2,

---

23 For example, suppose that degrees of irrationality are measured by the expected Brier score (see Staffel 2019: Chap. 4). Then, since it is *certain* that one of these ten propositions is true and the remaining nine are false, the first credence assignment has degree of irrationality 0.9 (= $1 \times 0.81 + 9 \times 0.01$), while the second credence assignment has degree of irrationality 2.5 (= $1 \times 0.25 + 9 \times 0.25$).

then having credence 0.1 in this true proposition does not have a very high degree of abstract propositional rationality; credence 0.1 is quite far from the ideally rational credence of 1. Nonetheless, this may be the credence that is yielded by the most rational available disposition, and in that sense it would be true in many contexts to describe this credence as "rationally held." The reason for this is that having credence 0.1 in this mathematical truth is still, as I put it in Section 13.2 above, a characteristic manifestation of a sufficiently rational disposition.

According to my account, it is only if the disposition manifested by an attitude is *ideally* rational—that is, if the disposition *infallibly* yields a *perfectly* rational attitude in *every* case in which it is manifested—that doxastic rationality guarantees propositional rationality. As I have argued, however, for thinkers like us such ideally rational dispositions are rarely if ever available. It is for this reason that we need an account that recognizes that both doxastic and propositional justification comes in degrees.

The "basing account," as I have argued, cannot provide the sort of account that we need. However, the rival "virtue manifestation" account that I have proposed here looks like a more promising alternative.

References

Audi, Robert (1983). "The Causal Structure of Indirect Justification", *Journal of Philosophy* 80 (7): 398–415.

Bird, Alexander (1998). "Dispositions and Antidotes", *Philosophical Quarterly* 48: 227–34.

Conee, Earl, and Feldman, Richard (1985). "Evidentialism", *Philosophical Studies* 48: 15–34.

———— (1998). "The Generality Problem for Reliabilism", *Philosophical Studies* 89 (1): 1–29.

Dogramaci, Sinan (2018). "Rational Credence Through Reasoning", *Philosophers' Imprint* 18 (11), 1–25.

Firth, Roderick (1978). "Are Epistemic Concepts Reducible to Ethical Concepts?", in Alvin Goldman and Jaegwon Kim, eds., *Values and Morals* (Dordrecht: D. Reidel): 215–29.

Glynn, Luke (2010). "Deterministic Chance", *British Journal for the Philosophy of Science* 61: 51–80.

Goldman, Alvin (1979). "What is Justified Belief?" in George Pappas, ed., *Justification and Knowledge* (Dordrecht: D. Reidel): 1–25.

Harman, Gilbert (1986). *Change in View* (Cambridge, Massachusetts: MIT Press).

Kahneman, Daniel (2011). *Thinking Fast and Slow* (New York: Farrar, Straus, and Giroux).

Lasonen-Aarnio, Maria (forthcoming). "Perspectives and Good Dispositions", *Philosophy and Phenomenological Research*.

Lehrer, Keith (1974). *Knowledge* (Oxford: Oxford University Press).

Millar, Alan (1991). *Reasons and Experience* (Oxford: Oxford University Press).

Neta, Ram (2019). "The Basing Relation", *Philosophical Review* 128 (2): 179–217.

Peacocke, Christopher (1986). *Thoughts: An Essay on Content* (Oxford: Blackwell).

Pollock, J. L., and Cruz, Joseph (1999). *Contemporary Theories of Knowledge*, 2nd edition (Lanham, Maryland: Rowman and Littlefield).

Siscoe, R. W. (2021). "Belief, Rational and Justified", *Mind* 130 (517), 59–83.

Smithies, Declan (2015). "Ideal Rationality and Logical Omniscience", *Synthese* 192 (9), 2769–93.

Sosa, Ernest (2007). *A Virtue Epistemology* (New York: Oxford University Press).

Staffel, Julia (2019). *Unsettled Thoughts: A Theory of Degrees of Rationality* (Oxford: Oxford University Press).

Tang, Weng Hong (2016). "Reliability Theories of Justified Credence", *Mind* 125 (497): 63–94.

Turri, John (2010). "On the Relationship between Propositional and Doxastic Justification", *Philosophy and Phenomenological Research* 80 (2), 312–26.

Wedgwood, Ralph (2002). "The Aim of Belief", *Philosophical Perspectives* 16: 267–97.

Wedgwood, Ralph (2007). *The Nature of Normativity* (Oxford: Oxford University Press).

Wedgwood, Ralph (2017). *The Value of Rationality* (Oxford: Oxford University Press).

Wedgwood, Ralph (2020). "The Internalist Virtue Theory of Knowledge", *Synthese* 197: 5357–78.

Williamson, Timothy (2000). *Knowledge and Its Limits* (Oxford: Oxford University Press).

Worsnip, Alex (2018). "The Conflict of Evidence and Coherence", *Philosophy and Phenomenological Research* 96 (1): 3–44.