Article:
Williams, JRG orcid.org/0000-0003-4831-2954 (2018) Rational Illogicality. Australasian Journal of Philosophy, 96 (1). pp. 127-141. ISSN 0004-8402

https://doi.org/10.1080/00048402.2017.1323933

RATIONAL ILLOGICALITY

J. Robert G. Williams

University of Leeds

Abstract

Many accounts of structural rationality give a special role to logic. This paper reviews the problem case of clear-eyed logical uncertainty. An account of rational norms on belief that does not give a special role to logic is developed: doxastic probabilism.

Keywords *logic, rationality, Bayesian, probabilism, uncertainty, accuracy*

1. The Puzzle

The liar sentence L says: L is not true. We reason:

1.	$L \vee \neg L$	(premise)
2.	if L, then $L \,\&\, \neg L$	(basic logic + the T-scheme)
3.	if $\neg L$, then $L \,\&\, \neg L$	(basic logic + the T-scheme)
4.	$L \,\&\, \neg L$.	(argument by cases from 1,2,3).

Rationality apparently requires that we be fully confident in tautologies like (1), that we utterly reject contradictions like (4), but that we do not allow our confidence to drop over a valid argument like the one from (1) to (4). These are incompatible demands. Something has to give.[1]

But what? Opinions vary, but I focus on two. Ada says the argument shows us that classical logic must be given up, and in particular, that (1) is no tautology. Rejecting (4), and mindful that we shouldn't have more confidence in the premise of a valid argument than in its conclusion, we must also reject (1). Beth says that we should continue to be fully confident in (1) and to fully reject (4), and somehow find fault with the argument. She notes that the reasoning essentially depends on the T-scheme—that L is true iff L. She rejects this assumption, and defends her view as a form of logical conservatism, by noting that this scheme is not strictly speaking part of classical logic as it appears in the textbooks.[2]

You listen to Ada and Beth's case for their positions, and see force on both sides. You are convinced that one is right, but are not sure which. You adopt the modest, reasonable hedged position on which you are agnostic about the propositions over which they disagree. In particular, while Ada utterly rejects $L \vee \neg L$ (thinking that it entails a contradiction) and Beth fully accepts it (thinking that it is a tautology), you have 0.5 confidence in it.

---

[1] Nothing particularly hangs on my choice of the Liar as a hook here. Almost any debate which involves competing philosophical (or indeed scientific) theories embedding rival logics can be used to make the point.

[2] Ada has been reading Field [2008], Beth has been reading Williamson [1998], though she's clearly ignored his more recent writings on semantic paradoxes and abductive methodology, where he attacks the idea of logic as a 'neutral arbiter'.

You are condemned from both sides. Beth and Ada agree that *either* rationality requires you to have full confidence in $L \vee \neg L$ or no confidence in it. They both regard your 0.5 confidence as violating a rational requirement. And since your view is that *one* of them gets things right, you agree that your modest hedged attitude is irrational.

2. How General Is the Problem?

What this threatens to establish is:

> **(Rationally Required Extremism)**
> Either rationality requires me to have credence 1 in $L \vee \neg L$, or it requires me to have credence 0 in $L \vee \neg L$.

Your middling credence violates a rational requirement, if this claim is true. Since (*ex hypothesi*) you believe this claim, by your own lights your attitude is irrational.

That it is *rationality* that is in play here is crucial. Recommendations for extremism are quite common. Let's suppose that Ada and Beth disagree about some contingent matter—Valentin's location. Ada thinks that he is in the kitchen; Beth thinks that he is in the garden. You think that one of them is correct but modestly hedge your attitude, having 0.5 credence in the kitchen option. Ada and Beth agree that one ought to have high credence in what's true and have low credence in what's false. Whence you accept:

> **(Doxastic Ought Extremism)**
> Either I ought to have high credence in Valentin being in the kitchen, or I ought to have low credence in Valentin being in the kitchen.

Your middling credence in Valentin being in the kitchen is not the attitude you ought to have, according to both Ada and Beth. You trust their joint verdict, so by your own lights your attitude is not as it ought to be. But here, there is no puzzle. The sense of 'ought' in which you ought to believe what is true is an externalistic, evaluative one.[3] We all (barring Descartes' pure inquirer) acknowledge that we are not believing exactly as we ought to, if truth irrespective of evidence is what sets the standards. We nevertheless maintain that our attitudes are the rational way to respond to the uncertainty we have about how we ought (in this external sense) to believe.

According to the picture that I work within, rationality provides the safety net norms that guide you in coping with uncertainty. Rational requirements are a matter of the proper internal patterning of your attitudes, not a matter of bringing them into correspondence with the world outside the mind.[4] Uncertainty about how you (externally) ought to believe is grist to the mill.[5] Uncertainty about rationality is special. If you agree with Ada and Beth that rationality requires an extremal credence in $L \vee \neg L$, you can't adopt a middling credence and present this as the rational way to respond to this uncertainty. For you have already conceded that rationality requires that you do not do this. You can no longer rely on the safety net to guide your attitudes

[3] This classification of 'oughts' is drawn from Florian Steinberger [2016].

[4] Believing as you ought to requires much more than being rational in the narrow sense, I assume. See Kolodny [2007] for an example of this separation of rationality and reasons for belief.

[5] Compare uncertainty about how you *morally-ought* to behave given an 'objective' understanding of moral-ought in the miner's puzzle [Regan 1980; Jackson 1991]. For related discussion of subjective and objective oughts in the moral sphere see Rosen [2003], Smith and Jackson [2006], Harman [2011], and Sepielli [2013].

and your further inquiry on the matter. If you knowingly adopted irrational attitudes in hedging your views on L ∨ ¬L, then you regard this attitude of yours as pathological; you should look for therapy rather than further evidence.

In sum: Ought Extremism in general is not by itself problematic, because in the general case we have rationality standing by to guide our responses to uncertainty about what to do or believe. But Rationally Required Extremism is problematic, because rationality can't guide you when the thing that you are uncertain about is what rationality's guidance is.

It is implausible that the modest, middling confidence you invest in L ∨ ¬L is a violation of a rational requirement. Indeed, given the balance of arguments, any *other* attitude seems unwarranted. I conclude that Rationally Required Extremism is false.

3. Rationality without Logic

My diagnosis is the following: Ada and Beth were wrong to hold that logic articulates requirements on rationality. It is not the case that you are rationally required to be fully confident of tautologies, to fully reject contradictions, and to preserve confidence over valid arguments. What is correct in the vicinity is the following:

> **(Doxastic Logical Ought Extremism)**
> Either (in light of the logical facts) I ought to have credence 1 in L ∨ ¬L, or (in light of the logical facts) I ought to have credence 0 in L ∨ ¬L.

This has exactly the same pattern as the ‘extremism’ about the correct attitude to have that derives from your uncertainty over Valentin's whereabouts. If this is all that logic gives us, we can still maintain that the rational response to our uncertainty about what the logical facts are is to have 0.5 confidence in L ∨ ¬L.

But we need to say more about what the rational constraints on belief are to be, if they're not ones that make appeal to logical facts. The idea of rationality as a safety net, guiding our responses to uncertainty, presupposes that there is some content to the guidance given. To defend rational illogicality requires spelling out what that guidance could be.

The agenda for the next few sections is as follows. I describe a familiar account of rational belief that incorporates logic in a way that would generate our original puzzle. I give what I regard as the best reason for thinking that the framework does generate rational requirements. Examining a dubious premise in that reasoning allows us to pinpoint a way of generalising the framework to allow for rational uncertainty about logic.

4. Bayesianism and Logic

Bayesians give a pleasingly concrete story about the rational requirements on belief—in particular, on partial belief.[6] As a background, let's take it that our agents have *degrees of belief* in those structured propositions that they adopt doxastic attitudes towards. ‘p(A)’ will denote the agent’s degree of belief in A. The Bayesianism in question requires these degrees of belief to be *probabilistic* in the following sense:

> **(Bayesianism—Structural Version)**

[6] Bradley [2015] is a recent user-friendly introduction to the Bayesian canon.

$p(A) = 1$ whenever A is a tautology.
$p(A) = 0$ whenever A is inconsistent.
$p(B)$ is no less than $p(A)$ whenever B is a logical consequence of A.
$p(A) + p(B) = p(C)$ whenever A and B are a logical partition of C.

This represents a standard form of Bayesianism, when you understand the logical properties invoked in the orthodox, classical way.

Some caveats and clarifications. First, I'm taking degrees of belief to attach to structured entities rather than, e.g., sets of events. That's familiar but not universal in this setting. Normally the structured entities are taken to be sentences, but construing them as structured propositions makes no formal difference. Second, if we wish to allow that agents can fail to adopt an attitude to some proposition, we need only make a modest tweak to the above: the rational requirement will be that there be some mapping from propositions to numbers p^*, coinciding with p where the latter is defined, such that p^* meets the constraints above. The agent's partial beliefs must be extendable to a probability function over a closure of the propositions on which the agent takes a stance.[7] Third, Bayesians differ about whether there are further rational constraints on belief over and above the ones mentioned, and most Bayesians supplement the static (beliefs-at-one-time) requirements above with dynamic (beliefs-over-time) requirements which concern how we update our credences upon learning new information. My discussion will be neutral on this front, focusing on the common core just identified.

If one favours a nonclassical logic, the characterization of Bayesianism just given still makes sense, though its content will be very different. Nonclassical Bayesianism has been explored in recent years by several authors.[8] For present purposes, the important thing is that the core of Bayesianism doesn't require allegiance to classical logic. But if we articulate it as above, it does require some logic or other. For that reason, I will call all these theories of rationality 'logical probabilisms'.

The upshot is this. Ada declares allegiance to a particular nonclassical logic (perhaps the Strong Kleene Logic), and articulates the rational requirements on belief via the corresponding nonclassical Bayesianism.[9] Beth declares allegiance to a classical logic, and so articulates the rational requirements on belief by her lights via standard classical Bayesianism. From the disjunction of these two ways of filling in the schematic version of Bayesianism above, we derive Rationally Required Extremism. A shared commitment to logical probabilism, therefore, produces our original puzzle.

5. Accuracy and the Subject's Point of View

There is extensive work in the Bayesian tradition that seeks to explain why probabilistic constraints should be taken to be rational requirements on belief. This will give guidance about how we might generalise the model to avoid our puzzle.

---

[7] In the unstructured setting, this would be the algebraic closure; in the structured setting, closure under logical operations.
[8] See Williams [2012a, 2012b, 2014a, 2016]. Particularly central precedents are Field [2000] and Paris [2001].
[9] Representing truth values as 1 (truth), 0 (falsity), and 0.5 (other), the Strong Kleene logic represents conjunction as minimum, disjunction as maximum, and negation as one minus negated value. An argument is valid iff it preserves value 1. See Williams [2016] for details and references.

The version of these results I like best is based on the accuracy arguments given by Jim Joyce [1998, 2009] and based ultimately on formal work by de Finetti [1974].[10] The starting point is the idea that we can rank individual degrees of belief by how 'accurate' or 'close to the truth' they are. For example, if it's true that Valentin is in the kitchen, then you're closer to the truth, or more accurate, the more confident you are in that very proposition. By aggregating this measure of accuracy over all the attitudes you adopt, we reach an overall evaluation of how accurate a belief state is, relative to a given possible world.

We then prove the following: if B (a function from propositions to real numbers) is improbabilistic, then there's an alternative function B*, such that for each *w*, B* is overall closer to the truth than B at *w*. So violating one of the constraints that we listed earlier means that you will be gratuitously inaccurate. There's an alternative set of attitudes you could have adopted that is guaranteed to bring you closer to the truth.

There's a lot to be said about this. For example: what exactly must accuracy be like, in order to be able to prove the result quoted?[11] But my interest is in the sort of grounding for rational requirements being offered here.

Here is the way I see that issue. We start with an externalistic, evaluative norm—that we ought to be as accurate as possible. That's comparable to an externalistic, evaluative norm such as: we ought to perform acts that in fact produce the most good. So in particular, when belief state B is overall more accurate than belief state B*, an agent *z* ought to believe B rather than B* (a contrastive verdict: given just those two options, B is the one to go for).

The comparative accuracy norm is exactly the kind of external norm that we know we're going to violate all the time. When Valentin is in the kitchen, the maximally accurate overall belief state will have me being fully confident in this. When he's not, the maximally accurate belief state involves full rejection of the same thing. I am certain one of these is the case. But lacking appropriate evidence I am 0.5 confident in this claim.

To get something that is attitude-guiding, we need to somehow introduce the subject's evidentially limited point of view. And the minimal way of doing this, I think, is that when the condition in question (for example, B being more accurate than B*, or act X producing more good than act Y) is by the subject's own lights met, then the subject really should be doing as the external norm recommends.

So the general pattern for 'subjectivising' external evaluative norms would be the following:

From Evaluative to Derivative Deontic Claims:
z **ought** X rather than Y when C(X,Y)
z **should** X rather than Y when, for *z*, *it must be* that C(X,Y).[12]

In the special case of maximizing the good we get:

[10] For discussion of analogous results using Dutch Books, see Williams [2012b].
[11] Pettigrew [2016] is a state of the art book-length discussion.
[12] The structure (and the switch between two deontic modals to mark it) is inspired by Williamson's discussion of the knowledge norm for assertion and belief in *Knowledge and its Limits* [2000], and its subjectivisation. My version appeals to doxastic necessity ('must'), rather than belief. Consider: I believe that pressing the button will get goodies at no cost to anyone. But it's possible by my lights that pressing brings despair and destruction. With appropriate credences, and sufficient despair, the thing I should do is not press. Unless belief excludes such doubt, the resulting principles will not be plausible.

z **ought** perform A rather than A* when A brings about better outcomes than A*.
z **should** perform A rather than A* when, for *z*, *it must be* that A brings about better outcomes than A*.

The subjectivised norm here represents a rather plausible dominance constraint on choice: when A is better than A* on every doxastic possibility, go for A!

In the case of accuracy, we get:

z **ought** have belief state B rather than B*, when B is more accurate than B*
z **should** have belief state B rather than B* when, for *z*, *it must be* that B is more accurate than B*

In the case of the evaluative/external norm of accuracy, the "subjectivised" version is a requirement to avoid a certain kind of accuracy-domination.

Not every subjectivisation of an external norm is a rational requirement. The claim that a given subjectivised norm (the one resulting from subjectivising accuracy) is a rational requirement I see as a theoretical identification, based on inspecting that norm and seeing that it is suited to play the rationality role, that is, that this norm will play the safety net role better than any competitor.

The above strategy, if it works, illuminates why certain conditions (anti-subjective-dominance constraints, in effect) have normative punch—why they are not just arbitrary rules defining membership of some club, but rather are intelligibly connected to the underlying values. Whether this connection is strong enough to really explain the normativity of the resulting rules is a nice question, and one pressed by Kolodny [2007] and Broome [2005] et al. in their work on reasons to be rational. But that the normatively binding rules are normatively binding rational requirements is a classificatory question orthogonal to whether they have normative force. On this point, more later.

6. Logical Versus Doxastic Possibilities

Take as a working primitive the notion of an agent's *doxastic space*, the set of worlds that are *doxastically possible* for her. Assume the following:

For *z*, it must be that *p* iff *p* is true at each *w* doxastically possible for *z*.[13]

This allows us to pinpoint a lacuna in the accuracy argument for probabilism just sketched. If one violates classical probabilistic constraints on belief, then there's some alternative belief state C that *across all logical possibilities* is more accurate than your starting point B. But the subjectivised version of the accuracy norm requires that we show that for the agent, it *must be* that C is more accurate than B. Unless we can assume that the doxastic possibilities are a subset of the logical possibilities, this does not follow.

In previous work [Williams 2012a] I extended accuracy arguments to the nonclassical case. But the same point goes through. The formal results I appealed to showed, for instance, that

[13] A world here is simply an arbitrary function from propositions to truth values. Characterizing the relation between doxastic possibilities (and the connected 'mights' and 'musts') and other notions (belief, knowledge, certainty) is a matter of substantive theory, not definitional analysis (though the principle that if *z* is certain that *p*, then for *z*, it must be that *p*, is very plausible).

if one violates nonclassical (say, Strong Kleene-based) probabilistic constraints on belief, then there's some alternative belief state that across all Strong Kleene possibilities is more accurate than your starting point. But that shows that the agent violates the subjectivised version of the accuracy norm only if the doxastic possibilities for the agent are a subset of the Strong Kleene possibilities.

Consider your plight when trying to hedge between Ada and Beth, and getting shouted at by both. You took the world-according-to-Ada as a doxastic possibility, and the world-according-to-Beth as another. But the world-according-to-Ada is a nonclassical possibility, so your doxastic possibilities are wider than the classical possibilities (even if the actual world is itself as classical as you like). Something analogous happens in reverse. The world-according-to-Beth contains a violation of the T-scheme, and since it is a doxastic possibility, your doxastic space is wider than Ada's nonclassical logical possibilities.

I conclude that the accuracy argument for probabilism is flawed if the target is any one of the variants of logical probabilism. The reason for this is independent of worries about whether logical probabilism is fit to play the safety net rationality role. The reason is rather that in order to use the accuracy argument to underpin probabilism as a constraint *whose normative punch is inherited from the evaluative norm of accuracy* we need to have some independent case that doxastic possibilities can't outstrip the logical possibilities. *Prima facie*, the needed extra premise is simply false, so the accuracy argument for probabilism fails.

7. Doxastic Probabilism

The accuracy argument for probabilism is flawed if it targets a logical probabilism. But it can support a different form of probabilism.

The general result on which accuracy arguments (and Dutch Book arguments, for that matter) rest is that if one's distribution of degrees of belief over propositions fails to take a particular form—if they aren't representable as expectations of the truth value of each proposition, relative to some underlying division of belief over possibilities—then bad things happen. In the case of accuracy, the bad thing that happens is that one's belief state will be accuracy-dominated by some alternative belief state across the underlying space of possibilities.[14]

To get the logical forms of probabilism, one needs to assume that the underlying space of possibilities contains all and only the logical possibilities. But I have argued that, to be philosophically cogent, the argument needed to start with a potentially wider space—the doxastic possibilities.

Logical probabilism is a thesis that concerns rational requirements on a single type of doxastic state—partial belief across propositions. But the true rational constraints, I contend, govern the agent's partial beliefs together with their modal attitudes---what is possible and impossible by their lights. *Doxastic probabilism*, as I will call it, requires that partial beliefs are representable as expectations of the truth value of each proposition, relative to some underlying assignment of credence (summing to credence 1) over the doxastic possibilities. The result of violating the constraints of doxastic probabilism will be accuracy-domination relative to the space of doxastic possibilities. This is exactly the condition that a subjectivised accuracy norm tells us to avoid.

[14] In the case of Dutch Books, the bad thing is a loss of money on a particular book of bets in each possibility in the underlying space.

The resulting constraints are still recognizably probabilistic. For example, if you regard w and u as the only two open possibilities, and p is true at w but not u, and q at u but not w, then if your beliefs are to be doxastically probabilistic, your degree of belief in p and q separately must sum to 1. (More generally, the function that describes your degrees of belief must be a convex combination of the functions that pick out the truth values of propositions at each world in your doxastic space, that is: if $t(X,w)$ is the truth value of X at w, then there are constants k_w summing to 1, such that for any A, $p(A) = \sum_w k_w t(A, w)$).[15] However, probabilistic constraints in this sense do not incorporate any logic.[16]

Doxastic probabilism is my answer to the central challenge for illogical rationality specified above—that of saying something substantive about what rationality requires, if we are not required to respect the true logic. In the following section, I illustrate this substance further.

8. Recognition Requirements and Conditionalized Norms

I set up the initial puzzle by assuming that the rational requirements allegedly associated with logic take an austere form. When C follows from A, it's required that you not have lower credence in C than in A. Austere formulations like this contrast with looser 'recognition' requirements on which the constraint on one's doxastic state only kicks in when (in some sense) you recognise or should recognise the logical fact in question. In Hartry Field's [2009] formulation, the constraints only kick in when it's *obvious* that C follows from A. In some of the formulations MacFarlane [2004] considers, the recognition element is that the subject *knows* the relevant logical property obtains.

Now recall your plight when attempting to hedge between the views of Ada and Beth. You are genuinely unsure whether L ∨ ¬L is a tautology. It's not obvious that it is; even if in fact it is a logical truth, you aren't in an epistemic position to know it. So even if Beth is correct that it is a tautology, if we work with recognition requirements rather than austere requirements, all acknowledge that you are not required to fully believe L ∨ ¬L. Likewise, you are genuinely unsure whether L & ¬L follows from L ∨ ¬L (and you're not obviously wrong in being so). So even if Ada is correct that it does, if we work with recognition requirements, all acknowledge that they do not bind you to fully reject L ∨ ¬L. On the recognition requirement formulations, neither Ada nor Beth will condemn you for your agnosticism over L ∨ ¬L. Some might present this as an adequate response to the initial puzzle, avoiding the need develop new forms of probabilism.

I want to make two points in response. The first will be that the best recognition requirement proposal is a *special case* of doxastic probabilism. The second will be that the full elaboration in doxastic probabilism is necessary, because the sort of piggy-backing specification given above only captures a fraction of the relevant rational requirements.

Now, the kind of recognition involved in plausible probabilistic recognition norms is going to be highly constrained. After all, it had better be that when you recognise that A is a tautology, you have degree of belief 1 that A; analogously for the recognition of consequences and contradictions. If you think that believing or knowing that A is a tautology is rationally

[15] Some Bayesians are committed to no more than this, and were always happy to include 'logically impossible' worlds in their underlying state space. Easwaran [2014] is one recent example.

[16] Caveat: in order that "convex combinations" of worlds may be defined in the usual sense, the truth values assigned to propositions need to be real-valued. See Williams [2016] for discussion.

compatible with not believing to degree 1 that A, then you cannot identify 'recognizing that A', in the sense in which it appears in the formulation of the norms, with these familiar attitudes.

'Recognising that A is a tautology' in the relevant sense, I suggest, entails that the subject is not doxastically open to any possibility on which A is not true. Likewise, when *z* recognises that C follows from A, *x* is not doxastically open to any possibility on which A is true and C is untrue; when *z* recognises that A is a contradiction, *z* is not doxastically open to any possibility on which A is true.[17] Perhaps recognition requires more than doxastic certainty in the sense just articulated, but it is at least as strong.

Compare this to the norms that doxastic probabilism predicts. For example, suppose an agent recognises that C follows from A. I've said that this entails that the agent has no doxastic possibility where A is true and C is untrue. But given doxastic probabilism, the agent's credences are constrained to be a probability function over her doxastic space, and so she will violate a rational requirement unless all credence devoted to worlds where A is true is credence devoted to worlds where C is true. Restating: she will violate a rational requirement unless the credence she devotes to C is at least as high as the credence she devotes to A. But that is the rational requirement the recognition norm was after—and it is predicted by doxastic probabilism plus the thesis that recognition (in the relevant sense) entails doxastic certainty. This observation generalises. Recognition norms are explained and predicted by doxastic probabilism.[18]

On the other hand, recognition requirements do not entail doxastic probabilism. First, there may be more to recognition than doxastic certainty: we might be doxastically certain that A is a tautology, even if A isn't obviously a tautology (maybe the proof took a long time, or perhaps was carried out by a friend or machine in whom we have complete confidence) and when we lack knowledge that A is a tautology (because we're convinced of it on the basis of a false logical theory, or have some false lemma in our reasoning, etc.). Second, recognition norms kick in only when we're doxastically certain of some fact, but these are just limit cases for doxastic probabilism. The latter but not the former continues to apply under conditions of logical uncertainty.

As a result, recognition norms alone undergenerate rational requirements. We saw that accuracy-dominance considerations afford an argument that one should conform to probabilistic norms whether or not one's doxastic space embodies *knowledge* of the correct logic, or whether the doxastic space is based only upon what is obvious. Especially when thinking of rational requirements as a matter of the proper internal patterning of mental states, I can't see why one would *deny* that doxastic certainty alone is sufficient to generate rational requirements---some specific consideration that the requirement lapses in the absence of extra factor X would be needed.

But the strongest case for going beyond recognition norms is their inability to handle cases of uncertainty. To illustrate, suppose that *modus ponens* is in fact valid for the English indicative conditional. But Vann McGee [1985] presents some *prima facie* plausible counterexamples (under our assumption, he's made a subtle mistake in concluding that they really are counterexamples). After reading the last few years of work on the semantics for conditionals, Sally has a good sense of what the truth values for the relevant propositions would be like, if McGee were right. She's pretty confident (to degree 0.8) that he's wrong, a level of

[17] To cover nonclassical cases, these principles need careful formulation, but appropriate formulations are available.
[18] In Field's [2009] version of recognition requirements, we will appeal to a normative entailment between obviousness and doxastic certainty, rather than a strict entailment. It is still doxastic probabilism that it is doing the heavy lifting.

confidence too low to count as recognizing the validity of the rule—it's certainly doxastically possible for her that *modus ponens* fails.

Suppose Sally is certain that A, but only 50/50 in C. She then becomes convinced that *if A then C* is true. Unless she adjusts her other beliefs (in A, C) or acceptances (0.8 in *modus ponens*), I take it that there's a clear violation of rationality in her attitudes. Her attitude to *modus ponens* means that she must, to degree at least 0.8, believe C conditional on things in which she is certain. The small uncertainty could excuse her being 0.8 in C, but there's no room for her to drop to 0.5. Theories of rational requirements need to account for the datum that Sally's attitudes are not rational. However, if all we had to go on were recognition norms, then we couldn't say why this would be. By construction Sally does not recognise the validity of *modus ponens*—her level of confidence is not high enough. So the triggering condition of the *modus ponens* recognition requirement isn't met.

Doxastic probabilism explains what goes wrong here. Let L be the set of worlds that cohere with a certain set of logical principles—for definiteness let them be the classical logical possibilities, with the indicative conditional interpreted as material conditional. Relative to L (and its subspaces) the doxastic probabilistic requirements include standard classical logical probabilism. This has relevance even to someone like Sally who only has 0.8 credence that some world in L is actual, because those classical probabilistic constraints constrain Sally's conditional credences, in particular, her conditional belief in beliefs that such-and-such is the case under the supposition L.[19] For example, Sally's credence in A given L, and in ¬A given L, must sum to 1. Essentially, that section of her belief state which she devotes to classical possibilities must behave like a classical probability. Further, there are doxastic probabilistic norms that constrain the relation between a subject's categorical belief in a proposition A and their conditional credences in *A given L*. Sally's categorical beliefs reported above mean that, on pain of violating doxastic probabilism's requirements, she will have degree of belief 1 in A given L and the same degree of belief in *if A then C* given L. On the other hand she can be no more than than 0.625 confident in C given L, given she is only 0.5 confident in C unconditionally. That pattern of beliefs violates the classical probabilistic norms on Sally's degrees of belief under the supposition that classical logic holds—and as argued above, this is a violation of a rational requirement of doxastic probabilism.

Doxastic probabilism doesn't just deliver recognition norms. It delivers unconditional norms (the old logical norms, in fact) on certain two-place attitudes, the agent's conditional credences. Since it also underpins a general analogue of the law of total probability connecting categorical and conditional belief, this constrains how agents who are uncertain about logic divide their belief over those very issues they are uncertain about. This is the phenomenon that underpins the application of doxastic probabilism to the McGee case, but it is entirely general. Recognition norms alone fall silent when we are not doxastically certain of the logical facts; but that means that are silent most of the time. They should be viewed as stepping stones, subsumed in the more general account presented here.

9. Incomplete Subjectivisation

[19] The rational conditional belief in A given L will match the conditional probability of A given L. That is, it equals the total credence invested in L-worlds in which A obtains, normalised by the total credence invested in L-worlds.

The last two sections address worries about my case for doxastic probabilism. In this section I consider whether the account is undercut by failures of epistemic transparency. In the next, I will consider whether it is subject to 'revenge' puzzles.

Recall that probabilism is a constraint that results from subjectivising the external evaluative accuracy norm. The pattern was as follows:

> *z* **ought** X rather than Y when C(X,Y)
> *z* **should** X rather than Y when, for *z*, *it must be* that C(X,Y).

The agent is irrational, I suggested, when a belief state other than her own is more accurate at every one of the doxastic possibilities—and in virtue of instantiating the above schema, she should not be in irrational states.

Doxastic irrationality has normative force in virtue of being a subjectivisation of the objective demand for accuracy. But the subjectivisation is not total—in particular, there's no guarantee that an agent who is irrational will be in a position to recognise this fact. Let me chart some ways in which this can arise.

Firstly, there's nothing in the above scheme that requires the agent to appreciate the normative significance of C, in order for the external or subjectivised rules to be in place. Imagine an agent who thought that the external evaluative norm for degrees of belief was their conduciveness to making money, rather than accuracy. They might deny that there's anything normatively defective about an (accuracy-dominated) set of beliefs they hold, simply because they regard accuracy domination (as opposed to financial domination) as irrelevant.

Secondly, an irrational agent may lack introspective access to the facts that make them irrational. An agent might appreciate that beliefs B accuracy-dominated beliefs B* across possibilities D, and fail to recognise that they themselves believe B* and are open to D.

Thirdly, an irrational agent might have a fix on their own beliefs and value accuracy, and not be in a position to justifiably work out that there's something defective with her beliefs. Cognitive limitations are the boring way this can happen, but more interestingly, the agent may not be able to work it out even in principle. The proof that there is an accuracy-dominating belief state B for each improbabilistic starting state B* is nonconstructive.[20] Suppose that classical logic is in fact correct, but that Sally is open to both intuitionistic and classical varieties of mathematics—her doxastic space features both kinds of possibility. It may be true that B accuracy-dominates B* across Sally's belief space (we can suppose B meets intuitionistic probabilistic constraints, but B* violates them). Yet there's no guarantee that she is in a position to figure out that this is the case, or even to see that there is any such B*— the proof that allows us to see this is a non-constructive one that she regards as dubious.

Sally doesn't seem to do anything wrong if she has belief state B*, when she's not in a position to come to justifiably believe that it is accuracy dominated across her doxastic space. By her own lights her pattern of mental states is in order. Sure, her values may be out of line with the true epistemic norms; her beliefs about her beliefs may be out of line with the facts, or the methods of proof she accepts out of line with the truly valid methods of proof. But such lack of alignment to correct opinion appears to be the external evaluations that generated the unproblematic extremisms in the initial section. The charge of irrationality was distinctively supposed to concern internal patterning of mental states, not matching up to external standards.

[20] The non-constructivity enters, for example, with an appeal to Brouwer's fixed point theorem.

Further: if we *are* open to incorporating external standards into our theory of irrationality at the end of the day, why not do so at right at the beginning, and convict (say) Ada and the one who is agnostic between Ada and Beth of irrationality, while being careful to emphasise that this does not bring with it a charge that she has messed up by her own lights?

These challenges raise questions about the point of having a theory of rationality in the first place. My view is the following: rational agents are those who can, and have reason to, use a certain set of tools for dealing with uncertainty, for extending their beliefs to previously unconsidered cases, for updating on new information, for deciding what to do, and so forth.

To illustrate: why is it crazy to take nomically necessarily false beliefs to be *ipso facto* irrational? Not because such beliefs are flawless—after all they're inaccurate! It's because our standard tools for handling uncertainty and inquiry into contingent matters (e.g. whether the trains are on time on this line) extend in the same way to handling uncertainty and inquiry into whether water is or is not H_2O, or whether things of equal mass accelerate under gravity at the same rate. You could of course label those who are anything other than certain of nomic necessities 'irrational', but this would deprive the category of interest from the point of view of the appropriate epistemic structures.

The same goes, I contend, for the proposal to take logically necessarily false beliefs to be *ipso facto* irrational. One who hedges on whether $L \vee \neg L$ seems *prima facie* rational because there seems no obstacle to her treating the case just as she does uncertainty over ordinary contingent or merely nomically necessary questions . She is reacting to the conflict in testimony between Ada and Beth just in the same old familiar way she reacts to other conflicts in testimony. Appearances could of course be deceptive: it might be that under examination it turns out that the standard (probabilistic) tools we use to handle ordinary uncertainty have no extension to this case. But doxastic probabilism vindicates the initial seemings here—it shows us how a standard suite of tools (probabilistic synchronic norms on belief, conditionalisation as a means of updating, standard rational choice) are generalised to cover the case.[21] So the theoretical joints cut at doxastic probabilism, not at logical probabilism. That's the principled case for marking the boundary as I do. It does not rely on a commitment to any kind of access-internalism about rationality.

The first challenge was to say why it is that, if Sally doesn't appreciate that she's messed up, we should classify her as irrational. I say: because if her attitudes do not fall into the right pattern, then she can't use the usual tools for dealing with uncertainty. That's a fact about her, whether she appreciates it or not. The second challenge is to say why we couldn't play this card earlier. I say: one could do so, but it would be a mere terminological victory. We would need a new term, rationality*, to mark the limits of applicability of the usual tools for dealing with uncertainty.

10. Revenge?

I started by considered two characters, Ada and Beth, who adopted two different positions on the rationally required credence in a certain proposition. That gave us Rationally Required Extremism, which was unacceptable. Doxastic probabilism dissolves that worry. But now I've given some substance to doxastic probabilism, can't we find some other matter for Ada and Beth to disagree about, which will manifest in a disagreement about whether to go along with what doxastic probabilism requires, or to adopt some other attitude? (i) A 'hedged' attitude between

[21] The resources are to be found in [Williams 2016].

the recommended positions will not be rational, by the lights of doxastic probabilism; but (ii) it will be intuitively rational, manifesting an admirable modesty in the presence of serious theoretical disagreement. This is a template for a revenge puzzle. But I have been unable to find compelling instances of this template. Perhaps the reader can do better: in this section I forewarn them of slips it is easy to make.

To fix ideas, suppose Chara is a doxastic probabilist, whereas Diana believes in some appropriately general form of Dempster-Shafer belief function theory. When both agree that there is no relevant evidence for or against a particular proposition, say, whether Valentin is in the kitchen, V. Chara regards an attitude of 0.5 confidence as rationally required, or at least will insist that our confidences in this proposition and its negation sum to 1; whereas Diana will say that in the acknowledged absence of evidence we should invest no confidence in this proposition, and none in its negation. [22]

Chara and Diana certainly disagree on epistemological theory; and certainly it seems rational, in the right evidential circumstances, to hedge between them on the question of what norms are in force. But this is not an instance of the puzzle, because no case has been made for (i). Doxastic probabilism requires we divide credence in a particular way among whatever doxastic possibilities we are open to. Derivatively, it requires we divide credence appropriately among propositions true at such possibilities. However, it does not constrain what propositions (normative or non-normative) are true at a doxastic possibility. So if the normative proposition *that doxastic probabilism fails* (DP) is true at a world that is possible for you, you can simultaneously be doing what DP requires and yet be doubting DP itself. In sum: doxastic probabilism does not rule out rational doubt over doxastic probabilism. (Perhaps a critic will see this very feature as objectionable, but that would be a whole different line of attack, and an attack on a feature that generalised probabilism shares with logical probabilism).

Perhaps Chara and Diana also adopt different attitudes to the proposition V. Chara has a 0.5 credence; Diana, one might think, adopts the distinctive Dempster-Shafer attitude: degree of belief 0 in V, and the same in ¬V. The cautious agent wishes to hedge between them—at any rate, do something other than immediately side with Chara against Diana. But assuming Chara is correct about what doxastic probabilism requires, that would be to be irrational.

I say that 'one might think' that Diana possessed the distinctive Dempster-Shafer attitude to V. But that is contentious. Probabilists and Dempster-Shafer theorists don't disagree only about normative constraints on attitudes, they differ on the space of available attitudes itself. For probabilists, to have degree of belief 0 in a proposition is always a way of rejecting it; the Dempster-Shafer theorist thinks it can be a way of being uncertain. One or other is wrong. Either way, Chara and Diana can share a common psychological attitude to V—utter uncertainty—yet sincerely report themselves in different ways, as a reflex of their overarching theoretical disagreement. Hedging on V itself is not called for, since (whoever's right) they do not take different attitudes in the first place.

The original Ada-Beth case raised no such complexities, since it worked with a common space of point-like credal states, with two normative theories of the rational patterning of the mutually recognised attitude.

I have been unable to find a convincing revenge puzzle. Perhaps none exists. But suppose that we could find such an instance. Would we be entitled to conclude that doxastic probabilism is false? No: no more than we could conclude that logical probabilism was false, simply by

---

[22] The case is personally pressing for me: the two approaches to attitudes to indeterminate survival developed [Williams 2014a, 2014b] respectively instantiate an analogue of the Chara-Diana situation.

inspection of the original Ada/Beth case. The case against logical probabilism that has been given in this paper was not the provision of a single counterintuitive consequence (philosophers are skilled at biting bullets). The case against logical probabilism is that there is a rival—generalised probabilism—that subsumes its successes, has a principled motivation, and also gives a more satisfying account of case from which we started. Even if the elusive revenge puzzle could be found, the verdict on generalised probabilism would await the provision of a more general rival theory.

11. Conclusion

Doxastic probabilism articulates the rational constraints on partial belief and doxastic possibility. Logic has no special role, though it usefully spells out general and tractable constraints doxastic probabilism places on interesting kinds of conditional credences. The account has substance; for example, it tells us about the rational constraints on those who are almost but not quite accept *modus ponens*, when they believe that p and that q if p. It subsumes and goes beyond 'recognition norms' that might be used to evade the puzzle cases. Finally, I've argued that there's a principled methodology underlying our discussions, and that we have as yet no reason to think it is threatened by revenge puzzles.[23]

[23] I presented versions of this paper in Bergen, Uppsala, Leeds, Reading, Princeton, Munich, Birmingham, Oxford and St Andrews. I am very grateful for all the very helpful comments and suggestions I received on those occasions, and from two referees for this journal. Thanks to Gail Leckie with whom I taught a graduate course on this material, and especially to Thomas Brouwer, whose line-by-line feedback was invaluable.

References

Bradley, Darren 2015. *A Critical Introduction to Formal Epistemology*, London: Bloomsbury.
Broome, John 2005. Does Rationality Give Us Reasons? *Philosophical Issues* 15/1: 321–37.
De Finetti, Bruno 1974. *Theory of Probability Vol.1*, New York: Wiley & Sons.
Dummett, Michael A. E. 1958. Truth, *Proceedings of the Aristotelian Society* 59/1: 141–62. Reprinted in Dummett 1978, *Truth and Other Enigmas*, London: Duckworth: 1–24.
Easwaran, Kenny 2014. Regularity and Hyperreal Credences, *Philosophical Review* 123/1: 1–41.
Field, Hartry H. 2000. Indeterminacy, Degree of Belief, and Excluded Middle, *Noûs* 34/1: 1–30. Reprinted in Field 2001, *Truth and the Absence of Fact*, Oxford: Oxford University Press: 278–311.
Field, Hartry H. 2008. *Saving Truth from Paradox*, Oxford: Oxford University Press.
Field, Hartry H. 2009. What is the Normative Role of Logic? *Proceedings of the Aristotelian Society* 83/3: 251–68.
Harman, Elizabeth 2011. Does Moral Ignorance Exculpate? *Ratio* 24/4: 443–68.
Jackson, Frank 1991. Decision-Theoretic Consequentialism and the Nearest and Dearest Objection, *Ethics* 101/3: 461–82.
Joyce, James M. 1998. A Non-Pragmatic Vindication of Probabilism, *Philosophy of Science* 65/4: 575–603.
Joyce, James M. 2009. Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief, in *Degrees of Belief*, ed. Franz Huber and Christoph Schmidt-Petri, Berlin: Springer Verlag: 263–97.
Kolodny, N. 2007. How Does Coherence Matter? *Proceedings of the Aristotelian Society* 107/1: 229–63.
MacFarlane, John G. 2004. In What Sense (If Any) Is Logic Normative for Thought? URL = <http://johnmacfarlane.net/normativity_of_logic.pdf>, accessed 18-08-2016.
Maher, Patrick 2002. Joyce's Argument for Probabilism, *Philosophy of Science* 69/1: 730–81.
McGee, Vann 1985. A Counterexample to Modus Ponens? *The Journal of Philosophy* 82/9: 462–71.
Paris, J.B. 2001. A Note on the Dutch Book Method, in *ISIPTA'01: Proceedings of the Second International Symposium on Imprecise Probabilities and their Applications, Ithaca, NY*, ed. G. de Cooman, T. Fine and T. Seidenfeld*,* Maastricht: Shaker: 301–6.
Pettigrew, Richard 2016. *Accuracy and the Laws of Credence*, Oxford: Oxford University Press.
Regan, Donald H. 1980. *Utilitarianism and Cooperation*, Oxford: Oxford University Press.
Rosen, Gideon 2003. Culpability and Ignorance. *Proceedings of the Aristotelian Society* 103/1: 61–84.
Sepielli, Andrew 2013. What to Do When You Don't Know What to Do. *Noûs* 48/3: 521–44.
Smith, Michael, & Jackson, Frank 2006. Absolutist Moral Theories and Uncertainty, *Journal of Philosophy* 103/6: 267–83.
Steinberger, Florian 2016. Explosion and the Normativity of Logic, *Mind* 125/498: 385-419.
Williams, J. Robert G. 2012a. Gradational Accuracy and Non-Classical Semantics, *Review of Symbolic Logic* 5/4: 513-537.
Williams, J. Robert G. 2012b. Generalized Probabilism: Dutch Books and Accuracy Domination, *Journal of Philosophical Logic* 41/5: 811-40.
Williams, J. Robert G. 2014a. Nonclassical Minds and Indeterminate Survival, *Philosophical Review* 123/4: 379-428.

Williams, J Robert G. 2014b. Decision Making under Indeterminacy, *Philosophers' Imprint* 14/4: 1-34.
Williams, J. Robert G. 2015. Logical Norms, Accuracy and Degree of Belief, in *Foundations of Logical Consequence*, ed. Colin Caret and Ole Hjortland, Oxford: Oxford University Press: pp.329-352.
Williams, J. Robert G. 2016. Non-Classical Logic and Probability, in *The Oxford Handbook of Probability and Philosophy,* ed. Alan Hájek and Christopher Hitchcock, Oxford: Oxford University Press: 248–276.
Williamson, Timothy 1998. Indefinite Extensibility, *Grazer Philosophische Studien* 55: 1–24.
Williamson, Timothy 2000. *Knowledge and its Limits*, Oxford: Oxford University Press.

Funding Information

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 312938.

University of Leeds