

Rationality in Hindsight
Aaron Bronfman

Policy debates frequently turn on competing predictions about the future. These may be predictions of the state of the world or of how the world is likely to be if certain proposed economic, social, or military policies are adopted. There is a way in which such disputes are often resolved with time: we observe the future and learn whether the predicted events occur. But there is another way in which such disputes are more resistant to resolution. Those whose predictions have been proven wrong may nevertheless insist that they were right in their evaluation of the evidence. They may claim that, although their predictions turned out to be wrong, it was in fact rational to make those predictions given the information available at the time. Such a position may, of course, be right. Yet it is not cost-free. I argue that, in a wide range of cases, the failure of a predicted outcome to occur constitutes some evidence that it was not rational to expect that outcome in the first place.

This paper proceeds in six sections. Section I describes a case of prediction and some initial grounds for doubting the thesis of this paper. Section II outlines the main argument. Sections III-V defend the premises of this argument. Section VI concludes.

I. Predictions and Outcomes

A. The Decision

The great (fictional) basketball player LeBron Jones has a decision to make. A free agent, he must decide whether to stay with his hometown Cleveland Cavaliers, or move to the up-and-coming Miami Heat. Where will he go? The nation awaits his answer, to be given tonight in a prime time special entitled *The Decision*.

A close follower of LeBron, you have read a lot about the factors that might affect his decision. Cleveland offers a higher salary and a longer contract, yet Miami offers a larger media market. The weaker Cleveland offers LeBron the chance to be its standout player, ensuring him the lion's share of credit for any championships, yet the stronger Miami offers a greater probability of winning a championship at all. Loyalty to hometown Cleveland and its devoted fans might weigh in its favor, yet tensions have been reported between LeBron and others on the team. Arguments rage over what to expect in light of recent decisions of other players and facts from LeBron's own biography.

After reflection on the evidence, you conclude there is no way to know where LeBron will go. Still, you are fairly confident that the evidence supports thinking it more likely that he will stay in Cleveland than that he will go to Miami, and so you expect him to stay in Cleveland. The case is complicated, though, so you allow some chance that the evidence supports thinking it at least as likely that LeBron will go as that he will stay.

Later that evening, you are surprised to hear LeBron announce that he will be moving to Miami. This outcome, you decide, will inform your predictions in future cases, perhaps leading you to give loyalty less weight than you had previously. However, it is not clear to you whether the

outcome shows that there was anything wrong with your initial assessment of the evidence. Although you will conduct your future reasoning differently now that you know what LeBron decided, that decision was not available information at the time when you made your prediction. Does the fact that LeBron chose Miami constitute any evidence that it was rational to expect him to choose Miami in the first place?

B. Hindsight Bias

A “yes” answer to the preceding question seems to endorse a form of “hindsight bias,” the phenomenon whereby “in hindsight, people consistently exaggerate what could have been anticipated in foresight” (Fischhoff 1980, 83). A “yes” answer endorses an apparently illegitimate use of information that is only available *ex post* to assess the reasonableness of an *ex ante* prediction. Indeed, a “yes” answer seems problematic for the following three reasons.

First, it is just not clear why learning the outcome would call for a revision in our assessment of what the initial evidence supported. It is a familiar fact that our evidence is sometimes misleading, in that it may lead us to reasonably expect something that does not in fact occur. Given this, it is not clear why we should not just write off the unexpected outcome as a consequence of misleading evidence, rather than use it as a basis for adjusting our initial assessment of what the evidence supported.

Second, the outcome can seem straightforwardly irrelevant to the question of what the initial evidence supported. The rationality of a prediction seems to be a matter of facts that are “in place” at the time the prediction is made. Indeed, if the rationality of a prediction is a matter of the agent’s evidence, then the agent seems to already be in possession of all the needed facts. From this point of view, to determine the rationality of a past prediction, the agent should simply attend to the evidence she had at the time; the outcome seems to be an irrelevant distraction.

Third, treating outcomes as evidence about *ex ante* rationality seems at odds with the idea that assessment of rationality ought to be based on a fair standard of assessment. In general, it seems inappropriate to criticize the rationality of a prediction on the basis of facts the agent could not possibly have known. Yet this is just what a critic would be doing in basing a criticism of an agent’s rationality on the actual outcome. The apparent unfairness of this maneuver suggests that the outcome is not evidence about the rationality of the prediction.

Contrary to the approach that classifies this as an instance of “hindsight bias,” I will argue that in cases like *The Decision*, outcomes are evidence about what it was antecedently rational to expect. When you learn LeBron actually decided to go to Miami, you have gained some evidence that your initial confidence that he would stay in Cleveland was not rational.

II. Outline of the Argument

A. Idea of the Argument

When hypotheses make predictions, whether absolute or probabilistic, the success or failure of these predictions provides evidence about the truth of the hypotheses. For example, consider two

hypotheses about the mechanisms underlying the Earth's climate. Suppose we know one or the other is true, but we are trying to figure out which. The two hypotheses differ in what they lead us to expect: the first hypothesis, if true, leads us to expect cooler global temperatures with high probability, and the second hypothesis, if true, leads us to expect warmer global temperatures with high probability. If we then observe warmer global temperatures, we have gained evidence against the first hypothesis and in favor of the second.¹

I will argue that hypotheses about what it is rational to believe are testable in just this way, because these hypotheses, too, make probabilistic predictions about the world. For example, consider two hypotheses regarding what it is rational to believe about whether LeBron will go to Miami. Suppose we know one hypothesis or the other is true, but we are trying to figure out which. The hypothesis that it is rational to be confident that LeBron will stay in Cleveland, if true, leads us to expect that he will stay in Cleveland. The hypothesis that it is rational to be confident that LeBron will go to Miami, if true, leads us to expect that he will go to Miami. If we then learn that LeBron chose Miami, we have gained evidence against the first hypothesis and in favor of the second.

B. The Argument

Before stating the argument, it will be helpful to identify the notion of rationality involved and to characterize the cases of interest.

The notion of rationality or reasonableness is substantive. It is a matter of what epistemic attitudes would be supported by excellent reasoning based on the agent's evidence. High-quality reasoning requires taking account of relevant evidence, noticing connections between pieces of evidence, relying on genuinely strong analogies between cases or between past and future, avoiding biases and formal or informal fallacies, etc. What the evidence supports, in this sense, may be a matter of both *a priori* plausibility and *a posteriori* considerations. Such a substantive notion of rationality seems to be at issue in numerous everyday arguments in personal and professional contexts, as well as in philosophical debates such as the debate over the epistemology of disagreement.²

There are two issues about rationality that I will largely put aside. One is the question of whether there is only one rational way to respond to a given body of evidence, or whether rationality permits a variety of possible responses to the same body of evidence.³ Another is the question of how precise the requirements of rationality are. For example, given a body of evidence, rationality might require a precise real-valued degree of credence in a given proposition, or it might require an imprecise or somewhat indeterminate attitude such as "high confidence."⁴

¹ In saying that we have (or have gained) evidence for a hypothesis, I do not mean to imply that our total evidence supports that hypothesis.

² See Kelly (2005) on the epistemology of disagreement.

³ See, e.g., White (2005) and Feldman (2007).

⁴ See, e.g., Levi (1980), van Fraassen (1990), Walley (1991), Joyce (2005), and Sturgeon (2008).

To avoid these issues, I will use relatively coarse-grained hypotheses about rationality, such as the hypothesis that *it is rational to be more confident of p than of $\sim p$* . This hypothesis is true when one's evidence permits one or more possible levels of confidence in p , all of which are determinately greater than one-half. For example, there may be a precise degree of confidence in p greater than one-half that is required. Alternatively, there may be a somewhat indeterminate level of confidence in p , which is nevertheless determinately greater than one-half, that is required. Or there may be a variety of these attitudes that are permitted, with no single one being required.

The cases of interest concern the question of what it is rational to believe about some proposition p , such as the proposition that LeBron will choose Cleveland. There is an initial time t_1 , at which one has evidence that bears on p but that does not conclusively establish or refute it. A key proposition of interest is the following:

R_p : At t_1 , it is rational for one to be more confident of p than of $\sim p$.

Two stipulations will help to simplify and focus our discussion. First, I stipulate that one knows with certainty that $\sim R_p$ holds if and only if, at t_1 , it is rational for one to be at least as confident of $\sim p$ as of p .⁵ Second, to narrow our focus to rationality itself, construed as a matter of the relationship between evidence and conclusion, I stipulate that one knows with certainty what one's evidence is. This ensures that any doubts one has about what one's evidence supports arise from uncertainty about the relationship of support between a body of evidence and a proposition, rather than uncertainty about the contents of the body of evidence itself.

After t_1 , there is a time t_2 , at which one learns whether or not p is true, without having gained or lost any other information. The argument aims to show that at t_2 , when one learns whether or not p , one has gained evidence about what level of confidence it was initially rational for one to have in p . As we will see below, there are two classes of cases (discussed in Sections III.B and V.A) that may warrant exceptions. Putting these aside, the argument goes as follows:

(P1) At t_1 , it is rational to be more confident that p on the assumption that R_p than on the assumption that $\sim R_p$.

(P2) In general, if it is rational to be more confident that p on the assumption that H than it is on the assumption that $\sim H$, then when one learns p , one gains evidence for H and against $\sim H$.

(P3) There is no special reason for doubting that the general connection in (P2) would apply to hypotheses about rationality.

(C) At t_2 , if one learns p , then one gains evidence for R_p and against $\sim R_p$.

(P1) expresses the idea that hypotheses about rationality make probabilistic predictions about the world. (P2) expresses the general idea that hypotheses are confirmed or disconfirmed as their probabilistic predictions are realized or not. (P3) applies this idea to hypotheses about rationality. In *The Decision*, this amounts to the conclusion that LeBron's decision to go to

⁵ Without this stipulation, $\sim R_p$ becomes a wide-ranging and fairly unwieldy hypothesis. For example, without the stipulation, $\sim R_p$ could hold because rationality puts no constraints whatsoever on what one may believe. In a more realistic case, these more exotic possibilities would not be ruled out entirely, but would perhaps warrant a small probability. The stipulation allows us to approximate such a case, while significantly simplifying the discussion.

Miami is evidence that, at the outset, it was rational to expect him to go to Miami, and not to stay in Cleveland.

III. Defense of (P1)

A. Argument for (P1)

(P1) At t_1 , it is rational to be more confident that p on the assumption that R_p than on the assumption that $\sim R_p$.

One's confidence in p on an assumption is understood as follows. In unconditional reasoning about whether p , one relies on one's evidence. In reasoning about p on an assumption, one relies on both one's evidence and the assumption. Thus in reasoning about whether the flowers will bloom, one bases one's opinion on evidence about the weather, the type of flowers, the time of year, etc. In reasoning about whether the flowers will bloom on the assumption that there will be a surge in the mouse population, one takes these additional hypothesized mice into account, considering what their presence indicates about the future of the garden in the context of the other factors listed.⁶

(P1) compares reasoning about p on the basis of two different assumptions. The comparison is between how confident one should be about p on the assumption that R_p and how confident one should be about p on the assumption that $\sim R_p$. (P1) asserts that one's confidence in p on the first assumption should be higher.

The argument for (P1) begins with our deference to experts. Broadly speaking, there are two features that help to qualify someone as an expert relative to oneself. First, an individual may have more relevant information than oneself. In this way, a child may be an expert on the characters in a television show. Second, an individual may have better relevant reasoning ability than oneself, even without having extra relevant information. An investment professional may be good at working out the likely consequences of a mixture of investments, even if she has no more information than an average person who happens to have memorized all the relevant data. An insightful friend may be good at reasoning about social relations and norms of etiquette, even without having more evidence about these things than you do. A talented student in medicine, engineering, law, mathematics, or science may also exhibit superior reasoning in her chosen field, even when compared with other students who have the same background.⁷

Let us narrow our focus to the second factor, considering experts who qualify as such, in a given case, relative to a given individual, solely because of their superior reasoning ability. Suppose the question is whether a federal agency's ruling is compatible with certain court decisions and laws. Although you have read all the relevant texts and formed a preliminary opinion, you are unsure of your capacities in legal reasoning. An expert lawyer, you suspect, will be better at seeing connections among the different laws, noticing subtle distinctions or shades of meaning, and keeping track of the important consequences. Imagine that such an expert has formed an

⁶ See Adams (1970) for clarification of the indicative mood in which these assumptions are made.

⁷ Hall (2004) calls the type of expert who possess superior information a "database-expert" and the type who has superior reasoning ability an "analyst-expert."

opinion on this question. If you do not know what her opinion is, you cannot defer to her simply by adopting her view. Even so, your deference to this expert should manifest itself in the following way. Your confidence that the ruling is compatible with the laws should be higher on the assumption that the expert is confident it is than on the assumption that the expert is *not* confident it is.⁸

(P1) claims that rationality itself merits our deference in this way. Consider the example just given. Instead of reasoning on the assumption that the expert lawyer is confident that the ruling is compatible with the laws, reason on the assumption that ideal reasoning on your evidence would lead you to be confident that the ruling is compatible with the laws. This is at least as weighty a factor as the lawyer's opinion. Indeed, it seems even weightier, insofar as the lawyer may occasionally miss a connection, blur a distinction, or lose track of a consequence. Ideal legal reasoning would make none of these errors. At the least, your confidence that the ruling is compatible with the laws should be higher on the assumption that it is rational to be confident it is than on the assumption that it is rational *not* to be confident that it is. (P1) simply agglomerates such case-specific judgments across all domains to say that this connection holds generally.

B. Discussion of Objections to (P1)

One objection to (P1) denies that we should ever defer to experts solely on grounds of their superior reasoning abilities. This objection would rely on the view that proper reasoning about a proposition *p* proceeds in a relatively straightforward fashion. Such reasoning relies only on "first-order" evidence about *p*. It does not involve consideration of "higher-order" evidence about what the first-order evidence supports. On this view, to reason well about whether the flowers will bloom, it is necessary and sufficient to appropriately weigh and combine evidence about the garden. Any other evidence, such as higher-order evidence about how facts about the garden should be weighed and combined, is (at best) superfluous. Of course, there is still a connection between proper reasoning and facts about one's evidence: because proper reasoning reaches the results that are supported by one's evidence, it *accords with* facts about what one's evidence supports. However, it cannot *appeal to* those facts as premises.⁹

To complete the objection, add the idea that the fundamental difference between unconditional reasoning and reasoning on an assumption is that in the latter case we add the assumption as a premise. Thus, when one reasons on the assumption that a superior reasoner who shares one's evidence is confident that *p*, one adds this assumption as a premise in one's reasoning. But if, as the objection holds, proper reasoning about *p* cannot appeal to facts about what one's evidence supports, then proper reasoning about *p* on this assumption is the same as proper reasoning about *p* without this assumption. Thus proper reasoning is wholly indifferent to assumptions about

⁸ I will not address the question of what import the opinion of an "epistemic peer," who generally reasons about as well as oneself, should have. Kelly (2005) distinguishes the two cases, expressing skepticism about whether the opinions of epistemic peers deserve weight, while claiming that "any plausible view" will give weight to the opinions of superior reasoners (174).

⁹ It is in the spirit of this view to allow that evidence bearing on what evidence other individuals possess may be used as a premise, though never evidence bearing on what one's own evidence supports. The view may make an exception for cases where proposition *p* is itself about what one's own evidence supports.

what one's evidence supports. Contrary to (P1), one's confidence in p should be the same on the assumption that R_p as on the assumption that $\sim R_p$.¹⁰

The main difficulty for this view comes from cases where it seems clear that evidence about what one's evidence supports does affect what it is rational to believe. Cases where one learns that a large number of superior reasoners have independently come to the same conclusion on one's evidence might suffice to establish this, but since this is roughly what is being called into question, I will use another kind of case. In this kind of case, one gains evidence that one's reasoning abilities may be impaired. It seems that this is relevant to what one should believe, even if one's reasoning has not in fact been impaired.

Consider a drug that has a tendency to interfere with subjects' reasoning about flowers. You have been given the drug in the past, and in 99% of trials, it has impaired your reasoning ability to the point where you draw absurd conclusions on your evidence. In each case where you are affected, you find these absurd conclusions quite plausible and notice no impairment. Imagine now that you have just been engaged in some reasoning about flowers and have confidently come to some conclusion. But then you learn that just before you began that reasoning, you were surreptitiously given the drug. It seems clear that your confidence in your conclusion should decrease, even if before you learned you were given the drug you were in fact reasoning perfectly on your evidence.¹¹

The upshot of this example is that one's confidence in a conclusion is not appropriately determined in the simple way suggested by the objection. For one's confidence that the flowers will bloom to be justified, it is not enough to weigh facts about the garden in just the right way. Instead, one must also take into account information about the rationality of possible ways of thinking, when such information is available. Because assumptions about what it is rational to believe posit just this type of information, the objection does not seem to provide a strong reason for doubting (P1).¹²

A second objection to (P1) concedes that one's confidence in p on the assumption that R_p should differ from one's confidence in p on the assumption that $\sim R_p$. However, it holds that the relation may be the opposite of that posited by (P1): sometimes, one should be *less* confident of p on the assumption that R_p than on the assumption that $\sim R_p$.

¹⁰ Christensen (2007b) considers a related objection, attributing the inspiration for it to Kelly (2005). The point itself cannot be attributed to Kelly (2005), in part because that paper holds that the views of superior reasoners *can* appropriately figure in one's reasoning.

¹¹ This case is based on similar cases in Christensen (2007a), (2007b), (2010), and (2011).

¹² Even so, it is not necessary to completely close off the objection, since a compromise position is available. On the compromise position, we concede that ideal reasoning would always consist of perfect first-order reasoning. As a rule, though, ordinary agents will not engage in this perfect reasoning, so they will be concerned with what kinds of reasoning are good, though less than ideal. The compromise position would maintain that part of such less-than-ideal reasoning consists in relying on the opinions of expert reasoners, when those opinions are known. This explains why we normally regard someone who defers to an expert reasoner as more rational than someone who does not. On the compromise position, (P1)-(C) would be reformulated in terms of this less-than-ideal standard of rationality. The revised (C) retains interest, since this less-than-ideal standard is prevalent in ordinary assessment of rationality. A similar position is discussed below in response to an objection to (P3).

I will give two cases to illustrate the possible problem. The first uses an explicitly self-referential proposition: let q be the proposition that it is not rational for you to be more confident of q than of $\sim q$. R_q is the proposition that it is rational for you to be more confident of q than of $\sim q$. By construction, q is materially equivalent to $\sim R_q$. Thus on the assumption that R_q , q is guaranteed to be false; and on the assumption that $\sim R_q$, q is guaranteed to be true. So it seems one should be far *less* confident of q on the assumption that R_q than on the assumption that $\sim R_q$.¹³ If so, then this is a counterexample to (P1).

A second case suggests the problem may generalize beyond quirky logical constructions. Suppose you are on a sinking ship with a single liferaft. Let S be the proposition that the liferaft will make it to shore. Although the liferaft has enough room to hold everyone, S is a bit less likely if the liferaft is nearly full than it is if the liferaft is nearly empty. Each passenger decides individually and simultaneously whether to get on board the liferaft. Each one also (i) has about the same evidence you do about S , (ii) has a tendency to hold opinions that her evidence supports, and (iii) is likely to get on board if confident of S but not otherwise. Let R_S be the proposition that it is rational for you to be confident of S . On the assumption that R_S , your evidence (and hence others' evidence) supports confidence in S , so it is likely that many passengers will be confident of S , and hence it is likely that many will get on board the liferaft; this, however, would tend to make S slightly less likely to be true. On the contrary assumption that $\sim R_S$, it is likely that few will get on board the liferaft; this would tend to make S slightly more likely to be true. Hence R_S has at least some negative evidential relevance to S : on the assumption that R_S , there is a reason for assigning lower confidence to S , while on the assumption that $\sim R_S$, there is a reason for assigning higher confidence to S . Perhaps this sort of effect, if strong enough, could make it rational, all things considered, to be less confident of S on the assumption that R_S than on the assumption that $\sim R_S$. If so, that would constitute a counterexample to (P1).

If the objection is correct, the situation is as follows. The assumptions R_p and $\sim R_p$ may have some special evidential bearing on whether p , beyond the bearing they possess in virtue of the appropriateness of treating rationality as an expert. The source of this special evidential bearing can vary. In the first case, it derives from logical features of the proposition. In the second case, it derives from tendencies of rational agents. The import of the special evidential bearing can also vary. In the first case, it is enough to entirely defeat the reasons for deferring to rationality as an expert. In the second case, it trades off as a weighing reason against the reasons to defer to rationality as an expert, and the ultimate outcome depends on the relative strength of the two factors.

As foreshadowed in Section II.B, because of the possibility of this kind of counterexample I will not argue that (P1) is true without exception. The troubling cases would be ones where the special evidential bearing is present, works against rather than in favor of (P1), and is strong enough to outweigh or entirely defeat the reasons to defer to rationality as an expert. In any particular application of (P1), we should be on the lookout for the possibility that these

¹³ This is similar to an example regarding objective chance that is given, with thanks to Frank Arntzenius, in Vranas (2004); it extends the idea of an “undermining future” in Lewis (1994). Related examples for practical rationality are considered by Gaifman (1983) and (1999), who argues that they motivate a hierarchy of levels of rationality, a possibility which I do not explore here.

conditions obtain. However, they do not seem to obtain in *The Decision* or in the majority of cases where we might want to assess rationality in hindsight. So I will take (P1) as established, on the understanding that we may need to allow an exception for the type of case just mentioned.

IV. Defense of (P2)

(P2) In general, if it is rational to be more confident that p on the assumption that H than it is on the assumption that $\sim H$, then when one learns p , one gains evidence for H and against $\sim H$.

An example will illustrate the application of (P2). Consider a doctor, about to observe a patient who may have the flu. Suppose that fever is quite common with the flu, but only occasionally seen in patients without the flu. In advance of observing the patient, it is rational for the doctor to be more confident that the patient has a fever on the assumption that the patient has the flu than it is on the assumption that the patient does not have the flu. The general rule described by (P2) applies to this case. It says, quite plausibly, that when the doctor observes the patient, and learns that the patient has a fever, then the doctor has gained evidence that the patient has the flu and evidence against the claim that the patient does not have the flu.

A full defense of (P2) will not be attempted here. Instead, I appeal to the intuitive plausibility of (P2), as a general principle or in particular cases, and refer the reader to two systematic theories of evidence that take (P2) as a central principle.

On the intuitive plausibility of (P2), it is worth noting that principles very much like (P2) have been regarded as truisms. Thus Joyce puts the following principle in the category of “important truisms about evidential relationships and facts about scientific practice”:

“*Prediction Principle*. If a person is more confident in E conditional on H than conditional on $\sim H$, then E confirms H for her.” (Joyce 2004, 143)

Although Joyce’s principle refers to how confident one actually is, rather than how confident it would be rational for one to be, the core idea is the same as that in (P2). Similarly, Hawthorne writes:

“Clearly, when the evidence is more likely according to one hypothesis than according to an alternative, that should redound to the credit of the former hypothesis and the discredit of the latter.” (Hawthorne 2011, 333)

Hawthorne presents this as a basic intuition. It, too, expresses the same core idea as (P2).¹⁴

¹⁴ More context may be helpful in understanding Hawthorne’s principle. He writes, “all theories of confirmation rely on measures of how well various alternative hypotheses account for the evidence. Most contemporary confirmation theories employ probability functions to provide such a measure. They measure how well the evidence fits what the hypothesis *says* about the world in terms of *how likely* it is that the evidence would occur if the hypothesis were true. ... Clearly, when the evidence is more likely according to one hypothesis than according to an alternative, that should redound to the credit of the former hypothesis and the discredit of the latter” (Hawthorne 2011, 333).

Further support for (P2) comes from the fact that it plays a key role in two systematic theories of evidence. The first is the likelihood theory of evidence, whose central principle, the “law of likelihood” is a variant of (P2).¹⁵ Thus Royall (1997) informally explains the law of likelihood as follows: “if x is more probable under hypothesis A than under B , then the occurrence of x is evidence supporting A over B ” (5).¹⁶ (P2) is more cautious than Royall’s formulation, since (P2) only compares a hypothesis and its negation, while Royall’s formulation allows for comparisons of arbitrary pairs of hypotheses A and B . On the other hand, Royall’s formulation is perhaps more cautious than (P2) in that it declines to speak of evidence *for* A ; instead, it allows only the relativized notion of evidence *for* A *over* B . The difference is not important for our purposes, so those who prefer the relativized notion of evidence should feel free to replace (P2)’s “evidence for H and against $\sim H$ ” with “evidence for H over $\sim H$.”

The second systematic theory of evidence in which (P2) plays a key role is the Bayesian theory of evidence.¹⁷ Bayesians count p as incremental evidence for H just in case one’s credence in H conditional on p is higher than one’s unconditional credence in H . Given Bayes’ Theorem, and if none of the relevant probabilities are zero, this condition holds exactly when one’s credence in p conditional on H is higher than one’s credence in p conditional on $\sim H$. Hence within the Bayesian framework, (P2) is very nearly a restatement of the Bayesian criterion of incremental evidence.

Finally, it is not essential to my argument that the general connection posited by (P2) hold without exception. It suffices that we can take this connection to hold in contexts where there is no special reason to doubt it. Accordingly, the next section argues that there is no special reason to doubt this connection in the present context.

V. Defense of (P3)

(P3) There is no special reason for doubting that the general connection in (P2) would apply to hypotheses about rationality.

The defense of (P3) consists of considering and replying to five possible objections. Each objection presents a candidate reason for doubting that the general connection in (P2) would apply to hypotheses about rationality.

A. Certainty about the Requirements of Rationality

One attempt to block the application of (P2) to the case of rationality would rely on the built-in exception of the Bayesian theory of evidence. As the Bayesian theory is commonly developed, if the probability of a hypothesis H is 0 (or 1), then the probability of H conditional on any proposition whatsoever is also 0 (or 1), unless the conditional probability is simply undefined.

¹⁵ See Hacking (1965), Edwards (1972), and Royall (1997) for defense of this theory.

¹⁶ Royall’s explanation continues, “... and the strength of that evidence is determined by how much greater the probability is under A .” As the point of this paper is simply to argue that outcomes are evidence about rationality, and not to measure the strength of that evidence, we need not discuss this second aspect of the law of likelihood.

¹⁷ See Horwich (1982), Earman (1992), and Howson and Urbach (2005) for defense of this theory.

This leads the Bayesian theory to deny that any proposition can be incremental evidence for a hypothesis when that hypothesis has probability 0 or 1.

To apply this exception in a blanket way to the case of rationality, the objection would need to maintain that hypotheses about what it is rational for one to believe, or about what degrees of credence it is rational for one to have, always merit probability 0 or 1. Presumably, any such hypothesis which is false merits probability 0, and any which is true merits probability 1. On this view, rationality is demanding enough to require certainty about what the requirements of rationality are.

This objection faces difficulties similar to those faced by the first objection to (P1). Consider an agent who has evidence that she has been given a drug that interferes with her reasoning ability in a general way. It seems implausible to claim that she should nevertheless be certain of what rationality requires of her, since she knows that the reasoning that leads her to conclusions about what rationality requires of her may be impaired. It is an interesting question just how she ought to respond, but some attitude short of certainty seems appropriate. This problem for the objection is quite general, as we all possess inductive evidence of our own fallibility and, often, case-specific evidence of distraction, bias, or less-than-perfect mental clarity. So the built-in exception of the Bayesian theory of evidence does not seem to provide a difficulty for (P3).

Still, one might wonder whether there are some cases where we should give no credence to the possibility of a reasoning error. Perhaps there are some propositions so basic or self-evident that we ought to be certain that confidence in them is rational. The fact that every candidate for such a proposition has been the subject of philosophical controversy should give us pause. But we need not decide the issue. If it turns out that there are such propositions, we may count them as exceptions to the present argument, as noted in Section II.B. In the policy disputes that motivate this paper, and in cases like *The Decision*, the evidence is complex, and agents like us have relevant evidence of our own fallibility. In such cases, it seems appropriate to assign some credence to the possibility that one has not assessed the evidence correctly.¹⁸

B. Explanatory Connection

A second possible reason to doubt the application of (P2) to hypotheses about rationality would consist of two claims. First, it would maintain that one proposition can only be evidence for another when we have reason to believe that there is an explanatory connection between the two: it must be likely that one proposition explains the other, or else that there is a common explanation of both.¹⁹ Second, it would maintain that this requirement is not met in the case of hypotheses about rationality and facts about outcomes. The objector might concede that there is often an explanatory connection between what evidence one has and what the outcome is, since the facts that causally explain the outcome can, by causing signs or symptoms earlier on, also

¹⁸ One still attracted to the view that rationality always requires certainty about the requirements of rationality could pursue a compromise position, along the lines of footnote 12. Such a position distinguishes between ideal rationality (which requires certainty about the requirements of rationality) and a less-than-ideal standard of ordinary rational guidance and advice (which does not). The present argument might be reformulated to take advantage of this distinction, but I do not develop such an approach here.

¹⁹ This condition is based on Achinstein (2001).

causally explain why we have the evidence we do. However, the objector might point out that what is at issue is not what evidence one has, but rather what a given body of evidence *E* supports. It is less clear whether this fact could stand in an explanatory connection with the outcome.

I will not assess either of the two component claims of the objection individually, but instead will argue that the two claims are not both true. The two claims imply that *p* cannot be evidence that *E supports being confident that p*. However, they also imply that *E supports being confident that p* cannot be evidence that *p*. The problem for the objection is that this latter implication seems to fail in cases where we rely on the testimony of superior reasoners.

For example, suppose you are a beginning weather forecaster, uncertain of your ability to evaluate your complicated meteorological evidence bearing on whether it will rain tomorrow. Fortunately, you have a colleague in another state who is widely known for her reliability in discerning what meteorological evidence supports. You share your meteorological evidence *E* with her, and she reports that *E* supports being confident that it will rain. Let us suppose that your trust in her is sufficiently high that you take her word for it. If so, you might spell out your reasoning about whether it will rain as follows: “My meteorological evidence relevant to whether it will rain is *E*. As my colleague has informed me, *E* supports being confident that it will rain. Therefore, it will probably rain.” Here, you appropriately cite the fact that *E supports being confident that p* as a reason for being confident that *p*. At least in these circumstances, the fact about what *E* supports seems to be legitimate evidence that *p*.²⁰ Because the two claims of the objection would render this impossible, they should not both be accepted.

C. Misleading Initial Evidence

Section I mentioned three possible reasons for doubting that the outcome of an event could be evidence about what it was antecedently rational to expect. The remaining objections correspond to these three possible reasons.

Thus the third objection asks why we may not just write off an unexpected outcome as a consequence of misleading initial evidence. As above, let R_p be the hypothesis that one’s initial evidence supports relatively high confidence in *p*. There is no incompatibility between R_p and $\sim p$. Cases where both hold are simply cases where the initial evidence was misleading. Thus someone initially confident of R_p might maintain that $\sim p$ constitutes no counterevidence to R_p ; in all probability, it merely shows that a genuinely unlikely and unpredictable outcome occurred.

The framework of this paper is helpful in showing why this objection is mistaken. A simple example illustrates the problem. Suppose I have two coins; one is an ordinary, fair coin, and the other has heads on both sides. I will pick one of these coins, leaving you uncertain as to which one I have picked. I will then flip the coin ten times, letting you observe the outcome. Suppose I

²⁰ One might try distinguishing between reasons and evidence, granting that facts about evidential support can function as reasons in the way described, but denying that this suffices to qualify those facts as evidence. We may put this issue aside, as we can reformulate our central question to ask about the reasons, rather than the evidence, provided by the outcome. Thus we can ask, “Does the outcome provide a reason for being more or less confident that it was antecedently rational to expect the outcome to occur?”

do this, and the observed outcome consists of a sequence of ten heads. Even though this outcome is compatible with the fair coin hypothesis, it nevertheless provides evidence against that hypothesis. Importantly, this point holds even if we add to the example the stipulation that you are initially justified in being highly confident, though not quite certain, that the fair coin hypothesis is true. With this stipulation, even after observing the outcome, you may be justified in being confident that the fair coin hypothesis is true and that a genuinely unlikely outcome occurred. Even so, you must acknowledge the outcome as providing some counterevidence to the hypothesis.

Returning to the case of interest, we have a hypothesis R_p which makes a probabilistic prediction that we learn to be incorrect when we learn the outcome $\sim p$. The objection is correct that, at least typically, R_p is compatible with $\sim p$. The objection is also correct that, sometimes, even after learning $\sim p$ we may be justified in being confident that R_p is true and that a genuinely unlikely outcome occurred. However, the simple example just given shows that these features of a case do not prevent the outcome from providing some degree of counterevidence to the hypothesis. These features do not undermine the expectation that the success or failure of a hypothesis's probabilistic prediction will bear on the truth or falsity of the hypothesis. Thus the third objection does not falsify (P3).

D. Intuition of Irrelevance

A fourth possible consideration that might falsify (P3) was also mentioned at the outset. It is the fact that outcomes may seem intuitively irrelevant to the question of what the initial evidence supported. This would, at least to some extent, differentiate the present application of (P2) from uncontroversial cases where it seems clearly applicable.

In responding to this objection, we can grant that hypotheses about what one's initial evidence supported do not depend, for their truth, on outcomes. Nevertheless, those hypotheses make probabilistic predictions about outcomes. Given this, it seems plausible that the success or failure of those predictions would be relevant to assessing the truth of those hypotheses. Thus it may be that merely recasting the question as one of assessing hypotheses that make probabilistic predictions suffices to undermine the intuition of irrelevance. But in any case, another response is available.

Consider a case where two individuals claim to be experts on a given subject matter. Both share the same evidence we do, but claim to have superior reasoning abilities. One individual is in fact a genuine expert who reasons extremely well on this evidence (better than we do). The other individual is impersonating an expert and forms her opinions through relatively poor reasoning. At the outset, we do not know which is which. A list of some possible future events, within the area of expertise, is given to both of them. Each gives us predictions as to which of those events will or will not occur.

Suppose things play out as follows. Over a long sequence of events, we notice that the predictions made by one of these individuals turn out to be mostly accurate, while the predictions of the other individual do no better than would be expected by chance. Intuitively, as the process continues, we gain evidence that the individual making accurate predictions is the genuine

expert. Perhaps we could never attain knowledge of which one was the genuine expert. Poor reasoning, through luck, could turn out to be a reliable guide to the truth even though there is no antecedent justification for trusting it. For example, a superstitious theory about winning lottery numbers might be turn out to be reliable if the lottery has been rigged to make that very theory true, even though there was strong antecedent reason to believe the lottery was fair. Still, it is intuitively plausible that, as one individual builds a superior record, we gradually accumulate evidence that that individual is the genuine expert. In light of this gradual accumulation of evidence, it seems that even at the first tested prediction where one individual is right and the other is wrong, we have gained a bit of evidence pointing in one direction.

A situation where we are trying to decide what it was antecedently rational to predict can be seen as a variant of the above situation. Imagine we are in possession of two lists of predictions for a sequence of events. One list predicts p just in case it is rational to expect p given our evidence. The other list predicts p just in case a certain style of poor reasoning on our evidence would lead to the conclusion that p . We do not know which list is which. But if, over time, we find that one list makes more accurate predictions than the other, then we seem to gain evidence that that list was based on good reasoning. As in the previous situation, our evidence seems to accumulate over time, beginning with the first tested prediction where one list is right and the other is wrong.

Framed in this way, it is intuitively plausible that the outcome of an event can be evidence about what it was antecedently rational to expect. Thus, at the least, we may say that intuition does not unequivocally speak against the present application of (P2). Because the initial intuition of irrelevance is opposed by a contrary intuition of relevance, it does not provide a strong reason against (P3).

E. Intuition of Unfairness

A final consideration that might falsify (P3), also mentioned at the outset, is the idea that there is a connection between rationality and fair criticism. Claiming that an agent fell short of a standard of rationality seems to constitute a form of criticism. This criticism, it might be said, must be based solely on the evidence the agent had available at the time; otherwise, it unfairly takes advantage of information the agent did not have. A criticism based on the actual outcome violates this requirement of fairness and hence cannot be a legitimate criticism of an agent's rationality at all.

An analogy is helpful in replying to this objection. Suppose you and I are both struggling to eat healthier. At dinner together, the waiter offers us some lavish cherry cake for dessert. Against your own better judgment, you indulge in a slice. I refrain. Afterwards, I say, "You should have resisted that cake." I neglect to mention that I had a bad experience with cherries as a child and find cherry cake wholly unappetizing.

You may acknowledge that my criticism is correct. However, it seems unfair. There is an air of superiority about it that seems inappropriate, given that my resolve was not even tested on this occasion. Indeed, even if you know about my lack of appetite for cherries, politeness seems to

favor openly acknowledging that I might well have given in, too, had it not been for that childhood experience.²¹

Compare that case with a case of outcome-based criticism of a prediction. You and I share the same evidence on some topic. Based on this evidence, you become confident in some prediction about the future. We later find that this prediction does not come true. I then say, “You should have reasoned better.” I neglect to mention that a crucial part of my reason for thinking this is the fact that the prediction did not come true.

If you are not certain what the initial evidence supported, then you should acknowledge that my criticism may be correct. Nevertheless, it seems unfair. There is an air of superiority about it that seems inappropriate, given that my reasoning abilities were not subject to the same test as yours on this occasion. Indeed, given that my criticism is crucially based on the outcome, there is reason to doubt that I would have passed that test. A more conversationally appropriate way of putting my criticism would be: “I might well have reasoned the same way you did given the initial evidence, but the outcome helps us to see that there was a better way to reason.” This does not seem to be unfair.

VI. Conclusion

I have argued that, in cases like *The Decision*, the occurrence or non-occurrence of a predicted outcome is evidence about whether it was antecedently rational to make that prediction. I have not attempted to quantify the strength of this evidence, and so my argument certainly does not show that proponents of failed policies must admit that their predictions were based on poor reasoning. Nevertheless, proponents of such policies must acknowledge that the case for the rationality of their predictions has become weaker, for we now possess a new piece of evidence against it.

²¹ This may be an instance of a general suggestion made by Brown and Levinson (1987) that criticism is less face-threatening when coupled with an assurance that the critic “considers himself to be ‘of the same kind’” as the addressee (72). If that suggestion is correct, then politeness favors offering such an assurance, at least when one can do so honestly.

References

- Achinstein, Peter. *The Book of Evidence*. Oxford: Oxford University Press, 2001.
- Adams, Ernest W. "Subjunctive and Indicative Conditionals." *Foundations of Language* 6 (1970): 89-94.
- Brown, Penelope and Levinson, Stephen. *Politeness: Some Universals in Language Use*. Cambridge: Cambridge University Press, 1987.
- Christensen, David. "Epistemology of Disagreement: The Good News." *Philosophical Review* 116, no. 2 (2007a): 187-217.
- Christensen, David. "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." *Oxford Studies in Epistemology* 2 (2007b): 2-31.
- Christensen, David. "Higher-Order Evidence." *Philosophy and Phenomenological Research* 81, no. 1 (2010): 184-215.
- Christensen, David. "Disagreement, Question-Begging and Epistemic Self-Criticism." *Philosophers' Imprint* 11, no. 6 (2011): 1-22.
- Earman, John. *Bayes or Bust?* Cambridge: MIT Press, 1992.
- Edwards, A. W. F. *Likelihood*. Cambridge: Cambridge University Press, 1972.
- Feldman, Richard. "Reasonable Religious Disagreements." In *Philosophers without Gods: Meditations on Atheism and the Secular Life*, edited by Louise Antony, 194-214. Oxford: Oxford University Press, 2007.
- Fischhoff, Baruch. "For Those Condemned to Study the Past: Reflections on Historical Judgment." *New Directions for Methodology of Social and Behavioral Science* 4 (1980): 79-93.
- van Fraassen, Bas. "Figures in a Probability Landscape." In *Truth or Consequences*, edited by J. Michael Dunn and Anil Gupta, 345-356. Dordrecht: Kluwer Academic Publishers, 1990.
- Gaifman, Haim. "Paradoxes of Infinity and Self-Applications, I." *Erkenntnis* 20, no. 2 (1983) 131-155.
- Gaifman, Haim. "Self-Reference and the Acyclicity of Rational Choice." *Annals of Pure and Applied Logic* 96 (1999): 117-140.
- Hacking, Ian. *Logic of Statistical Inference*. Cambridge: Cambridge University Press, 1965.
- Hall, Ned. "Two Mistakes about Credence and Chance." *Australasian Journal of Philosophy* 82, no. 1 (2004): 93-111.

Hawthorne, James. "Confirmation Theory." In *Handbook of the Philosophy of Science, Volume 7: Philosophy of Statistics*, edited by Prasanta Bandyopadhyay and Malcolm Forster, 333-389. Amsterdam: Elsevier, 2011.

Horwich, Paul. *Probability and Evidence*. Cambridge: Cambridge University Press, 1982.

Howson, Colin and Urbach, Peter. *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court, 2005.

Joyce, James M. "Bayesianism." In *The Oxford Handbook of Rationality*, edited by Alfred Mele and Piers Rawling, 132-155. Oxford: Oxford University Press, 2004.

Joyce, James M. "How Probabilities Reflect Evidence." *Philosophical Perspectives* 19 (2005): 153-178.

Kelly, Thomas. "The Epistemic Significance of Disagreement." *Oxford Studies in Epistemology* 1 (2005): 167-196.

Levi, Isaac. *The Enterprise of Knowledge*. Cambridge: MIT Press, 1980.

Lewis, David. "Humean Supervenience Debugged." *Mind* 103, no. 412 (1994): 473-490.

Royall, Richard. *Statistical Evidence: A Likelihood Paradigm*. London: Chapman and Hall, 1997.

Sturgeon, Scott. "Reason and the Grain of Belief." *Noûs* 42, no. 1 (2008): 139-165.

Vranas, Peter. "Have Your Cake and Eat It Too: The Old Principal Principle Reconciled with the New." *Philosophy and Phenomenological Research* 69, no. 2 (2004): 368-382.

Walley, Peter. *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall, 1991.

White, Roger. "Epistemic Permissiveness." *Philosophical Perspectives* 19 (2005): 445-459.