Evidential Symmetry and Mushy Credence

Roger White

MIT

rog@mit.edu

draft 4/07

The Principle of Indifference (a. k. a. Principle of Insufficient Reason) instructs us to distribute our credence sharply and evenly over possibilities among which our evidence does not discriminate. Once thought to be *the* central principle of probabilistic reasoning by great thinkers like Laplace, the Principle of Indifference has fallen on hard times. These days it is commonly dismissed as an old-fashioned item of confusion.[1] While I certainly haven't found my way through all the difficulties with the principle, I want to suggest that we need to rethink the matter. I want to argue that the objections to the principle are not as devastating as they appear. And that there is a compelling case in support of the basic idea, which cannot be ignored. Lastly, I'll consider the most important alternative to the principle involving the idea one's credence should not always be sharp.

1. Evidential Symmetry

Let's say that

> Propositions *p* and *q* are **evidentially symmetrical** (I'll write this as $p \approx q$) for a subject if his evidence no more supports one than the other.

I mean to understand *evidence* very broadly here to encompass whatever we have to go on in forming an opinion about the matter. This can include non-empirical evidence or reasons, if there are such. We might say: *p* and *q* are evidentially symmetrical for you iff you have *no more reason* to suppose that *p* is true than that *q* is, or vice versa.

---

[1] Recent critics include van Fraassen (1989), Strevens (1998), Gillies (2000), North (ms.), and Sober (2003).

There are different ways that evidential symmetry can arise. I'm wondering if the marble taken from this urn is black or white. I might have a rich body of data relevant to each proposition that bears on them in the same way:

> **Case 1**: I know that the marble in the box has been selected from a shaken urn containing just five white and five black balls.

Or I might have no relevant evidence either way:

> **Case 2**: I have no idea how or why this marble was chosen, or the constitution of the urn from which it came (for all I know it might contain unequal numbers of white and black balls, but I have no evidence concerning which color it has more of, if any).

In each case I have no more reason to suppose that the ball is black than that it is white. In the first we might say that I know that the *objective chances* of the two hypotheses are equal. In the second I have no reason to prefer one answer over the other as I have no reasons bearing on the matter at all. And of course there are various intermediate cases involving, say, partial knowledge of chances. One question I'll be concerned with is what the epistemological significance of the difference between these cases is.

2. Principle of Indifference

The Principle of Indifference links evidential symmetry and rational credence in a way that is blind to the difference between the cases above.

Principle of Indifference POI: $p \approx q \rightarrow \mathrm{P}(p) = \mathrm{P}(q)$

Here P(.) is a subject's *rational subjective probability*, or *credence* function. Let me set aside a common misunderstanding to begin with. One often hears: "You can't get probabilities out of ignorance."[2] Let's be clear that the principle, as I am understanding it, puts a *normative* constraint on what your *credence* may be. It entails that in a position of ignorance you are not rationally permitted to be more confident in one proposition than another. It is not to be confused with a principle for determining what the *objective probabilities* or *chances* are. Obviously ignorance is no basis for a *belief* concerning what the chances are. But it is not at all out of the question that ignorance puts constraints on what attitudes are rational to take. I hope we would agree that if I have no more reason to suppose that it will rain than that it won't, then it would not be reasonable for me to *absolutely certain* that it will rain, or even fairly certain. The POI takes this idea further by insisting that if I am to be *any* more confident that it will rain than not, I had better have some *reason* for this difference of opinion. An obvious corollary that is often called POI is:

POI*: If $\{p_1, p_2, \ldots, p_n\}$ is a partition of your knowledge such that $p_1 \approx p_2 \approx \ldots \approx p_n$, then for all i $P(p_i) = 1/n$.[3]

3. Multiple partitions problem

The most famous objection to POI alleges that it leads to inconsistent conclusions. A probability space can be partitioned in different ways. If a proposition *p* is a member of two evidentially symmetric partitions S_1 and S_2 of different size, then POI gives inconsistent answers as to what your credence in *p* should be. The most compelling examples of the problem involve continuous parameters non-linearly related. Here is an example based on Bas van Fraassen's (1989) cube factory story.

Mystery Square: A mystery square is known only to be no more than *two* feet wide. Apart from this constraint, you have no relevant information concerning its dimensions.

---

[2] As Michael Strevens (1998) voices the concern, "It is surely the case that we can never get reliably from ignorance to truth…The fact that we do not know anything about A does constrain the way things are with A."

[3] I'm taking for granted here that one's credal state is best ideally represented by a single probability function mapping each propositions onto one real number. I'll be questioning this assumption later.

What is your credence that it is less than *one* foot wide?

It would appear that you have no more reason to suppose that the square less than 1 foot wide than that it is more than 1 foot wide. But its area could be anything from 0 to 4 square feet. Have we any more or less reason to suppose that it is less that 1 square foot, than that it is between 1 and 2 square feet, or between 2 and 3, or 3 and 4?

So we have two possible partitions:

L_1: $0 \leq$ length < 1 ft.	A_1: $0 \leq$ area < 1 sq. ft.
L_2: $1 \leq$ length ≤ 2 ft.	A_2: $1 \leq$ area < 2 sq. ft.
	A_3: $2 \leq$ area < 3 sq. ft.
	A_4: $3 \leq$ area ≤ 4 sq. ft.

There is no coherent way to assign probability evenly over both partitions. But the POI seems to entail that we should:

(1) $L_1 \approx L_2$	Premise
(2) $A_1 \approx A_2 \approx A_3 \approx A_4$	Premise
(3) $P(L_1) = 1/2$	(POI)
(4) $P(A_1) = 1/4$	(POI)
(5) $P(L_1) = P(A_1)$	(L_1 and A_1 are equivalent)

Contradiction.
We got to this contradiction by POI, so POI must be false.

That's not quite right. We also used two premises (1) and (2). Why suppose these are true? Well, it's hard to *see* any reason to suppose that L_1 is true rather than L_2, or to believe any A_i more than another. It appears that we have no reasons bearing on the matter at all.

Nevertheless, here is an argument to suggest that (1) and (2) are not both true.
With the help of the following two principles that seem obviously correct, we can derive an absurd conclusion from (1) and (2) without any use of POI.

Transitivity: If $p \approx q$, *and* $q \approx r$, then $p \approx r$.[4]

Equivalence: If p and q are known to be equivalent, then $p \approx q$.

(1) $L_1 \approx L_2$	Premise
(2) $A_1 \approx A_2 \approx A_3 \approx A_4$	Premise
(3) $L_1 \approx A_1$	(Equivalence)
(4) $L_2 \approx (A_2 \text{ v } A_3 \text{ v } A_4)$	(Equivalence)
(4) $A_2 \approx (A_2 \text{ v } A_3 \text{ v } A_4)$	(Transitivity)

Consequence: We have *no less reason* to suppose that the area lies between **1 and 2** sq.ft, than to suppose that it lies between **1 and 4** sq.ft.

But this seems obviously wrong. Surely we have at least some more reason to believe the logically weaker $(A_2 \text{ v } A_3 \text{ v } A_4)$ than to believe A_2. If somehow we could rule out A_3 and A_4, then $(A_2 \text{ v } A_3 \text{ v } A_4)$ and A_2 might be epistemically on a par. But A_3 and A_4 could easily be true even if A_2 is not. So if I'm to believe only what is true, $(A_2 \text{ v } A_3 \text{ v } A_4)$ is a safer bet than the more specific A_2. Since this odd conclusion follows from the premises without any use of POI, this casts doubt on the premises that were used to refute POI. This response generalizes to other versions of the Multiple Partitions objection.

This conclusion is in many ways unsatisfying. One is apt to ask,

> Alright, so what *should* my credence be that the square is no more than a foot wide, according to POI? If your argument is sound and so (1) and (2) are not both true, it follows that either we have more reason to believe one of $\{L_1, L_2\}$ over the other, or we have more reason to believe one of $\{A_1, A_2, A_3, A_4\}$ over others. So which is it then, and what is this elusive reason?

---

[4] Elliot Sober and Branden Fitelson have independently objected to Transitivity by appeal to sorites cases. I'm not yet convinced, although the matter requires more thought and can't be addressed properly here. In any case, the question in this context is whether we really think that my minimal application here involves a violation of transitivity and that this accounts for the odd conclusion.

Well, okay so I don't really have an answer. Part of what is puzzling here stems from the temptation to think that my reasons or evidence must be transparent to me. The arguments above suggest that I have some reason to believe one proposition over another, yet I have no clue as to what this reason might be, or even which proposition it supports. These reasons, if there are such, seem to be rather mysterious, accessible if at all only by enlightened souls. But it is tempting to think that reasons that are beyond my ken can't really function as reasons for *me*. These are murky waters that I can't wade through here. Suffice it to say that there are reasons to resist the temptation to think that your reasons or evidence must always be known to you (see e.g Williamson 2000).

That still leaves us with a different sort of complaint:

> Perhaps your argument casts doubt on the claim that the Multiple Partitions Problem proves that POI is *inconsistent*. But at least appears to be *useless*. If we can't tell when propositions are evidentially symmetric because reasons to believe can be so elusive, then we can't apply POI to determine what our credence ought to be.

Yes, I suppose POI is pretty useless when it comes to the mystery square and other such cases. It by no means follows that it is always useless. There are Bayesian principles of coherence and conditionalization that I'm just too dense to apply to some cases, either because they are too complex or too confusing. But that doesn't stop me usefully applying them in easier cases. Perhaps in plenty of cases I can tell perfectly well that various possibilities are evidentially symmetric.

4. Arguments for POI

4.1 Argument from Cases: In many cases we clearly should assign credence $1/n$ to each of n alternatives. E.g. there are three doors. Behind one is a prize. Before any doors are opened, what is your credence that the prize is behind door A? In textbook cases like this we have no trouble answering 1/3. Only POI can account for this.

Reply: While we might be required to distribute our credence evenly in the suggested cases, it is really because we have some implicit knowledge that is relevant to the probabilities, we have some kind of stochastic model that justifies this distribution. In such cases we should just apply the Principal Principle (Lewis 1980) and set our credence to the known chances. There may be other cases in which we have no such knowledge but are still tempted to think that an assignment of equal credence is called for. We should resist this temptation which might be based on a confusion with cases of legitimate application of known chances.

As you can imagine, tossing back and forth cases like this quickly leads to a stalemate But here is a case to think about that is based on part of the Sleeping Beauty problem (Elga 2000).

> **Sleeping Beauty Case:** You know that you will be awakened on Monday and again on Tuesday, but your memory of the first awakening will be erased before the second. Now you find yourself awake and unable to tell whether it is now Monday or Tuesday. How do you divide your credence?

Although just about everything that could be said has been said about the original Sleeping Beauty problem and the broader issue of *de se* credence and updating, as far as I know everyone thinks that in the case above your credence should be divided evenly at ½. (We can make the case more extreme: you are to be awakened every day for a year with no memory of previous awakenings. Surely you should put *low* credence in today being November 24th.) But in this case it is hard to make sense of some kind of stochastic model at work. There is no process whose outcome determines whether *today is Monday or Tuesday*. There is no chancy event like a coin flip such that if it turns out one way it will now be Monday and if it turns out the other it will now be Tuesday. If we can speak tenselessly, you *are* awake on Monday *and* awake on Tuesday (and you know it). Your uncertainty is just about whether it is *now* Monday or Tuesday. So it is hard to see how we can be appealing to any kind of implicit stochastic information in giving ½ credence to it's being Monday. Rather it seems just to be matter of our ignorance concerning what day it is. You have no more reason to suppose it is one day rather than the other.

4.2 Argument from Statistical Inference: Many who are hostile to POI see at as a spooky *a priori* thing where you mysteriously conjure probabilities out of pure ignorance. Many people who feel this way will contrast it with the case of basing one's credence on *known statistics*, good solid data concerning the frequency of certain types of event or attributes in a population. For example, you somehow learn that 37% of formal epistemologists are left-handed. With *only* this to go on, what should your credence be that Branden Fitelson is left-handed? The very natural answer is 37%. Of course the matter gets tricky if I also know that, say, 17% of UW Madison graduates are left-handed, but have no direct information about the frequency of left-handedness among formal epistemologists from Madison. Or I might have seen him catch a ball once with his right hand and twice with his left. But provided we have none of this kind of messy conflicting data, and nothing more specific to go on, setting one's credence to known frequencies seems clearly correct. The general principle at work here would seem to be something like this.

Frequency-Credence (FC): If (i) I know that a is an F, (ii) I know that $freq(G|F) = x$ (the proportion of Fs that are G), and (iii) I have no further evidence bearing on whether a is a G, then $P(a \text{ is a } G) = x$.

Perhaps there is a more general principle covering cases where I also know that a is an H and that $freq(G|H) = y \neq x$. It is notoriously difficult to come up with a satisfactory one. Still, something like the more restricted FC seems straightforward. Quite a number of philosophers who express doubts or even hostility to POI seem to endorse something like FC.[5] The curious thing is that FC entails POI*.

Proof: Let $F = \{p_1, p_2, \ldots, p_n\}$ be any set of disjoint and exhaustive possibilities, such that $p_1 \approx p_2 \approx \ldots \approx p_n$. Let G be the set of *true* propositions. (i) I know that p_1 is an F (i.e. that $p_1 \in \{p_1, p_2, \ldots, p_n\}$) (ii) I know that $freq(G \mid F) = 1/n$ (exactly one member of the partition $\{p_1, p_2, \ldots, p_n\}$ is true) and (iii) I have no further evidence bearing on whether p_1 is G (I am ignorant concerning the p_i, with no more reason to suppose that one is true rather than another).

---

[5] For example Hacking (1965), Hacking (2000), Hajék (2007), Kyburg (1977).

Hence by FC, $P(p_1$ is a $G) = 1/n$, i.e. $P(p_1$ is true$) = 1/n$, so $P(p_1) = 1/n$. The same holds for each of the p_i, so POI* follows.

It's an illusion that POI* involves magically getting knowledge out of ignorance while FC solidly grounds probability judgments in data. Of course this might cast more doubt on FC than support POI*. But since we obviously should use frequency data to guide our subjective probabilities in *something* like the FC way, we need to think about what is the correct principle in this vicinity.

It's worth noting here the parallel between the Multiple Partitions problem and a version of the Reference Class problem. Just as a proposition may be known to be a member of different partitions of a space of possibilities, an individual like Branden can be known to be a member of different classes. Naïve applications of either POI or FC can lead to inconsistent results. And we lack an adequate story concerning how to extend either principle to the tricky cases. But now while this is a difficult problem, it is seldom taken to show that frequency data are *never* a legitimate guide to credences. Should the Multiple Partitions problem be taken to show that POI has *no* legitimate application?

Nevertheless, quite independently of worries about POI, some philosophers such as Isaac Levi (1977) have objected to FC saying that we are required to match credence to frequency only in cases where we have *randomly selected* the item in question from the population. It is not just that we must not have reason to suppose that the way in which we came to consider Branden as our example was biased in favor of (or against) left-handers. We must know that the process by which he was selected was an objectively random one, i.e., each individual in the reference class had equal objective chance of being selected. According to this account, learning that Marble A is among a hundred marbles in certain in a certain urn, 37 of which are black, is of little relevance in assigning credence to A's being black. But suppose we shake the urn and select a marble (it might happen to be A) in a suitably random manner. Then prior to seeing its color our credence that it is black should be 37%.

It is a little unclear to me how the appropriate kind of random selection is meant to work in less contrived cases. I ride a motorcycle and know something of the accident statistics for riders in different classes. Surely this sort of information should inform my credence that I will have a crash. (The insurance companies are certainly using it!) What would it mean for me to be randomly selected in the relevant way? It is not as if I picked someone at random out of the directory of motorcyclists and it happened to be me. I started with myself and went looking for statistical data that might apply to me. So I have trouble seeing how this random selection condition on FC is supposed to be applied if it is to do justice to our actual inferential practices.[6]

But furthermore, if we really need to do this random sampling then we can just do it. Let's take all the formal epistemologists and put them in a big urn. We will shake it vigorously and pick someone out. It happens to be Branden. Are we now supposed to have 37% credence that he is left-handed? If so then even if we don't get Branden the first time, we could just keep randomly selecting until we do. While this might be fun, it doesn't seem to be necessary.

4.3 Evidentialist Argument: One's confidence should adequately reflect one's evidence (or lack of it). You need a good reason to give more credence to *p* than to *q*. Hence if one's evidence is symmetrical so should be one's degrees of confidence. This is the fundamental thought behind POI and it can't be easily dismissed. What other option is there? Could there be some other rule determining what your credence should be in cases of evidential symmetry, particularly in those cases where the symmetry is due to ignorance? Perhaps the rule is that one's credence must be divided in the ratio .327 : .673 among two possibilities. This is silly. To even be applicable we must specify some difference between the two propositions to determine which one gets to enjoy the .673 and which the .327. And the only factor that could sensibly play such a role in determining these credence assignments would

---

[6] The motorcycle case might not be so straightforward as the frequency information might be relevant to estimates of *chances* of accidents in various situations, which chances are then a guide to my credence. Perhaps most ordinary cases have an element of this. But I don't think that's essential. Suppose an oracle who is known to see the future reveals that 99% of motorcyclists will have an accident this summer. She assures me that this is *not* because the conditions will be any more dangerous than usual. There will just be a fluky series of unlikely accidents (much as a perfectly fair coin will occasionally have a long run of heads). Shouldn't I have a high expectation of an accident this summer (given, of course, that this won't stop me riding)?

be some kind of evidence or reasons in support of one proposition. Alternatively we might hold a *permissive* view according to which no particular credence distribution is rationally required; any of some range is a rational option.[7] How wide is this range? How about 1:0? No, it would be nutty to be certain that *p* rather than *q* on the basis of no relevant evidence. (Talk about getting knowledge magically out of ignorance!) Would .9 : .1 be okay? That's just about as bad for the same reason. Smaller divergences from a nice .5 : .5 might not be *as* crazy as being highly confident that *p* rather than *q* for no reason. But the same scruples that prevent us from the more drastic imbalances of opinion reveal that ideally we should just split our credence evenly unless there is a reason to do otherwise.

Reply: There is something right and something wrong about this line of thought. Yes, evidential symmetry demands symmetry of opinion. But by failing to distinguish cases of *known chances* and cases of *ignorance*, the follower of POI *fails* to adequately represent his evidential state. The mistake is to suppose that one's opinion must always be represented by the simple model of a single standard probability function. Only in cases of known equal chances should one's credence be divided sharply and evenly in this manner. Ignorance calls for a different kind of state in which one's credence is spread, as it were, over a range of values. (Joyce 2005)

I find this the most compelling alternative to the standard POI. The idea that one's credence should in some sense cover range of values seems to be orthodoxy these days. The phenomenon is given a number of different names, such as 'indefinite' credence (Joyce 2005), 'vague probability' (van Fraassen 1990), 'imprecise probability' (Walley 1991), 'thick confidence' (Sturgeon 2008), and others. I've been following Elga (ms.) in calling it 'mushy' credence, although the novelty of this terminology is already wearing off. It is perhaps not entirely clear that everyone has the same idea in mind. But there is at least a common kind of formal model whose main elements are these:

1. **Representors**: A rational subject's state of opinion is best represented by a *set* of probability functions called his **representor *R***.

---

[7] I discuss permissive epistemologies more generally in my 2005.

2. **Updating**: Each of the functions in the representor is to be updated by conditionalization.

Often we are just interested in the spread of values in our representor for a particular proposition, i.e. the range of values for which there is some function in one's representor assigning that probability to the proposition. I will speak of one's *credence* in a proposition as being possibly a set of values. To avoid confusion I'll use 'P' for standard probability functions mapping propositions onto single real numbers, and 'C' for one's credence function as follows: $C(p) = \{x : \exists\, P \in R,\ P(p) = x\}$.

The key idea behind this response to the evidentialist argument for POI is the following:

> **Chance Grounding Thesis**: Only on the basis of known chances can one legitimately have sharp credences. Otherwise one's spread of credence should cover the range of possible chance hypotheses left open by your evidence.

We can illustrate the idea with the following case from Joyce 2005. I have before me three urns that you know to be taken from a collection $\{urn_0, urn_1, \ldots, urn_{10}\}$ such that urn_i contains i black balls and $10 - i$ white balls. Here is what you know about each urn.

> ***U_1***. You know only that this is urn_5.
>
> ***U_2***. You know only that this urn was chosen at random from the eleven urns.
>
> ***U_3***. You know nothing about which urn this one is, or how it was chosen.

In each case, what credence should you have that a random selection from the urn yields a *black* ball? For U_1 the answer is clearly ½. According to Joyce, since you no relevant chance information concerning *U3*, your credence that it will yield a black ball should be [0,1]. As for U_2, you don't know how many black and white balls it contains, so you don't know this urn's chance of yielding a black ball. But you have some higher-order chance information. You know that U_2 was selected from the urns such that each urn had the same chance of being selected. So you should have sharp credence $C(U_2 = urn_i) = 1/11$ for all i. And sharp conditional credence $C(black_2 \mid U_2 = urn_i) = i$. So your credence that U_2 will yield a black ball should be $C(black_2) = \sum_i C(black_2 \mid U_2 = urn_i)\, C(U_2 = urn_i) = ½$.

Joyce's example illustrates the idea behind the Chance Grounding Thesis. But it also raises a puzzle related to our earlier discussion of random selection. Suppose all eleven urns are lined up in a row in an unknown order. You have no idea by what method they were arranged. With respect to each of these urns you would appear to be in the same situation as with U_3 above. You have no relevant information concerning chances. So on Joyces account it seems that your credence for each urn that it will yield a black ball should be mushy over the range [0,1]. I might point to each urn in order along the row asking 'What's your credence that this one will yield a black ball?' And in each case you shrug your shoulders, unable to pin it down to any range narrower than [0,1].

But now I start jumping around flinging my arms around wildly over the urns. You happen to know that my arms are tossing about in an objectively random manner, so that there is an equal chance of my hand landing on any particular urn at a time. Each time I touch an urn I ask once again, 'What's your credence that this one will yield a black ball?' Now you situation seems similar to U_2 above. You know that this urn that I direct your attention to has bee selected at random from the eleven. So on Joyces account your credence that it will yield a black ball should be sharply ½. Somehow by magically waving my hands over the urns I have sharpened up your credence dramatically. This can't be right.

5. The Coin Puzzle

The following case illustrates something extremely puzzling about the mushy credence picture as a response to evidential symmetry.

Coin game: You haven't a clue as to whether *p*. But you know that I know whether *p*. I agree to write '*p*' on one side of a fair coin, and '~*p*' on the other, *with whichever one is true going on the Heads side.* (I paint over the coin so that you can't see which sides are *heads* and *tails*). We toss the coin and observe that it happens to land on '*p*'.

Let $\boldsymbol{C}$ and $\boldsymbol{C}_+$ be your rational credence functions before and after you see the coin land, respectively. The following five propositions are jointly inconsistent:

(1) $C(p) = [0, 1]$

(2) $C(heads) = [½]$

(3) $C_+(p) = C_+(heads)$

(4) $C_+(p) = C(p)$

(5) $C_+(heads) = C(heads)$

I'll consider each in turn.

(1) $C(p)$ = [0,1] Prior to seeing the coin land your confidence in *p* should be spread over the entire [0,1] interval in a maximally agnostic state.

According to the mushy credence response to POI, this is the attitude one should take toward a proposition in a position of complete ignorance. It is the most noncommittal attitude one can take (or what I once heard someone describe as 'the ultimate shrug'). Let's suppose this is so for *reductio*. I think the problems raised here arise for less extreme mushiness also. But for simplicity we can focus on the extreme case.

(2) $C(heads)$ = [½] Prior to seeing the coin land, your confidence that it will land *heads* should be sharply ½. In other words, *every* function in your representor should assign ½ to *heads*.

Your credence should be set to the known objective chance of *heads*. This is an uncontroversial application of what Lewis (1980) called the Principal Principle. And at any rate this principle is just intended to accommodate obvious facts like (2).

(3) $C_+(p) = C_+(heads)$ Upon seeing that the coin lands '*p*', your confidence in the propositions *p* and *heads* should be the same.

You know that '*p*' is on the *heads* side iff *p* is true, and that it landed on '*p*'. So you know that it landed *heads* iff *p*. Hence your attitude to these propositions should be the same.

If we accept the assumptions thus far, some change in my attitude to either *p* or to *heads* is required upon seeing the coin land. Either my credence in *p* should *sharpen* to match *heads* at

½, or my credence in *heads* should *dilate* to match *p* at [0,1].[8] The last two premises say that no such change is appropriate.

(4) $C_+(p) = C(p)$ Seeing the coin land '*p*' should have no affect on your credence in *p*.

We might deny this and say that upon seeing the coin land '*p*' you should sharpen your credence in *p* to ½. But note that if we say this then for symmetrical reasons we will have to say the same if the coin lands '~*p*' instead. For if the coin lands '~*p*' you will then know that *p* is true iff the coin lands *tails*, and your credence in tails will be ½. On this view you will know in advance that your credence in *p* will be ½ *no matter how the coin lands*. But this can't be right. If you really know this in advance of the toss, why should you wait for the toss in order to set your credence in *p* to ½?

I won't spend so much time addressing the possible denial of (4) as many of the arguments I make in the next section for (5) can be modified to apply to (4). But more importantly, the suggestion that one should sharpen one's credence to ½ in this case is incompatible with the standard rule of updating each function in one's representor by conditionalization. The following is a theorem of probability.

Irrelevance: For any probability function *P*, $P(p|e) = P(p|\sim e) \rightarrow P(p|e) = P(p)$

If I am to update by conditionalization such that all of the fuctions in my representor converge to ½ when I see the coin land (whether it lands '*p*' or '~*p*'), then prior to the toss each function must be such that

(*) $P(p|\textit{coin lands 'p'}) = P(p|\sim\textit{coin lands 'p'}) = 1/2$

[8] I suppose someone could say that my credence in both *p* and *heads* should change and meet in the middle at some narrower range of values. I don't think this suggestion requires separate consideration.

But by Irrelevance if follows that for each P in my representor, $P(p) = ½$, which contradicts our assumption (1) $C(p) = [0,1]$.

(5) $C_+(heads) = C(heads)$ Seeing the coin land 'p' should have no effect on your credence in *heads*.

Our last option to save (1) is to deny (5) and say that your credence in *heads* should *dilate* to [0, 1]. As before, if we take this option then for symmetrical reasons we must say the same thing if the coin lands '$\sim p$'. This option is actually mandated by the standard account of updating. For when we learn that the coin landed 'p' we learn that ($p \leftrightarrow heads$); if the coin lands '$\sim p$' then we learn ($\sim p \leftrightarrow heads$). Now, each function in my representor must be such that

(**) $P(heads \mid p \leftrightarrow heads) = P(p)$, and

(***) $P(heads \mid \sim p \leftrightarrow heads) = 1 - P(p)$

Proof of (**):

$$
\begin{aligned}
P(heads \mid p \leftrightarrow heads) &= P[heads \,\&\, (p \leftrightarrow heads)]/P(p \leftrightarrow heads) \\
&= P(p \,\&\, heads)/P[(p \,\&\, heads) \vee (\sim p \,\&\, \sim heads)] \\
&= P(p \,\&\, heads)/[P(p \,\&\, heads) + P(\sim p \,\&\, \sim heads)] \\
&= P(p)\, P(heads)/[P(p)\mathrm{P}(heads) + P(\sim p)P(\sim heads)] \\
&= P(p)/[P(p) + P(\sim p)] \\
&= P(p)
\end{aligned}
$$

(Similar proof for (***))

According to (1), for all $x \in [0,1]$, I have a function P in my representor such that $P(p) = x$. It follows now from (**) and (***) that when I update on either ($p \leftrightarrow heads$) or ($\sim p \leftrightarrow heads$) (i.e. on the coin's landing 'p' or its landing '$\sim p$') I will have a P in my representor such that $P(p) = x$, for all $x \in [0,1]$. That is, my credence in *heads* dilates from $C(heads) = [½]$ to $C_+(heads) = [0,1]$.

Having noodled about this puzzle on and off for some time, I discovered that the general phenomenon of dilation is old news.[9] Some statisticians and philosophers have studied how the phenomenon arises in other cases and appear to have taken it in their stride. This is not a *reductio* but a *result*, they might say.[10] I want to suggest that the present case brings out particularly forcefully how bizarre this phenomenon is. Here are a series of objections to the view that one should dilate one's credence in *heads* to [0,1] upon seeing the coin land one way or another.

Objection 1. Known Chance: You still know that the coin is fair. You have lost no information. How could the information you've gained have any relevance to your attitude to heads? Lewis (1980) was careful to point out that once you gain some 'inadmissible' evidence concerning the outcome of a fair coin toss, you might legitimately have credence other than ½ in the coin's landing heads. Suppose that in the present case instead of being clueless as to whether *p* you had some reason to suppose that *p*. In this case upon seeing the coin land '*p*' you would have gained some evidence that the coin landed *heads*. For now your evidence suggests that the coin landed on the side with the *true* proposition, and you know that the true proposition is on the heads side. (Note that in this case we don't get the same result if the coin lands '~*p*'; if the coin lands '~*p*' you've gained some evidence that the coin did *not* land *heads*). So it is not out of the question that seeing the coin land '*p*' could be relevant to you credence in *heads*. But in the original case *you haven't a clue as to whether p*. You have *nothing* to suggest that the coin landed heads or that it landed tails. Shouldn't you just ignore this useless bit of information and keep your credence in *heads* at ½?

Objection 2. Reflection: It is natural to suppose that if you know that you will soon take doxastic attitude *A* to *heads* as a result of *rationally* responding to new information *without loss of information*, then you should *now* take attitude *A* to *heads*. (This is a generalization of Bas van Fraassen's (1984) Reflection principle). Why would you need to wait until you actually gained this information to change your attitude? It would be different if you expected to lose your marbles tomorrow and have poor judgment. Or perhaps forget some important fact that

[9] See for example Seidenfeld & Wasserman 1993, and Walley 1991.

[10] Others like van Fraassen (2005, 2006) find the phenomenon more disturbing.

would make a difference to your conclusion (Talbot 1991). But as long you know that you will be epistemically fine then you should trust your future judgment and match your current attitude to it. But now according to the current proposal you know in advance that once you see the coin land you will rationally dilate your credence in *heads*. So you ought to just dilate it now before the coin even lands. But this is to deny (2) which is absurd.

We can make a similar point considering the opinions of another. Suppose you know that Scott takes attitude *A* to *heads*. You also take him to (a) have exactly the same 'priors' as you, (b) be impeccably rational, and (c) to possess *all of your knowledge plus more*. Surely you should trust his judgment and adopt attitude *A* yourself. That is after all the attitude that *you* would take if you were in his shoes, and he has the epistemic advantage of having more information to go on. This is just an easy case of how we trust experts more generally. But now suppose the coin has been tossed and you know that Scott has seen it land while you have not. According to the current proposal, without even having to ask him you know that Scott's credence in *heads* is now mushy (for he is as ignorant as to whether *p* as you). So now yours should be mushy too. We have reached the absurd conclusion that your credence that this fair coin lands *heads* changes from ½, not by receiving any information about how the coin landed (not even whether it landed '*p*') but just by learning that a rational onlooker has learned whether it landed '*p*'.

Objection 3. Mushy Betting: The state of having credence $C(p) = [0,1]$ can seem rather similar to $C(p) = [½]$. In each case you are no more inclined to suppose that *p* than to suppose that ~*p*. If there is an important difference between the two states of opinion then presumably it will be manifest in the behavior of a rational agent. If your credence in *p* is $[x, y]$, at what odds should you bet? There are many possible answers, but here are the two most common answers I've heard:

Liberal: Take any set of bets that maximizes expected utility according to some credence function in your representor.

Conservative : Only bets up to $x : 1\text{-}x$ on *p*, or up to $1\text{-}y : y$ on ~*p* are permissible. Decline a bet offered in between if you can.

These prescriptions are different enough that I suspect they reveal rather different understandings of what it is to be in a mushy credal state. But let's consider them in turn.

Objection 3.1. Liberal betting: We are supposing that C_+(*heads*) = [0,1]. So upon seeing the coin land one way or another, among other things you have function *P* in your representor such that *P*(*heads*) = ¾. At this point someone offers you a bet at 2:1 on *heads*, (i.e. if it landed *heads* you win $1; if not you lose $2). According to the Liberal account you can, as it were, plump for *P*(*heads*) = ¾ as your credence for betting purposes. And you maximize utility according to this function by taking the 2:1 bet. So when you see the coin land you can bet at odds 2:1 on *heads*.

But that seems mad. Let's repeat the whole game over and over using a series of propositions $p_1, p_2,\ldots,$ such that $C(p_i)$ = [0,1]. On each toss when you see the coin land either 'p_i' or '$\sim p_i$' you will dilate your credence that the coin landed *heads* on that toss to $C_+(heads_i)$ = [0,1]. And by the Liberal account you will be permitted to bet at 2:1 on $heads_i$. I think we know what will (almost certainly) happen when you do this. This *fair* coin we are using will land *heads* about half the time and sooner or later you will go broke.

Here is another Liberal betting worry that is related to Reflection. Suppose you are offered a 2:1 on *heads* prior to the coin toss. According to the dilation story, prior to seeing the coin land you know that you will soon have credence $C_+(heads_i)$ = [0,1] when you seen the coin land. And if you accept Liberal betting, then you know that it will soon be rational for you to take 2:1 bets on *heads* (if offered only that bet or nothing). But in that case prior to the coin toss you can rationally adopt a policy to accept such a bet once you have seen the coin land. But adopting such a policy is the same as just betting on *heads* at 2:1 now, *prior to the toss*. But that's mad. Your credence in *heads* prior to the toss is sharply ½. You can't sensibly take a bet at 2:1 in that case.

Objection 3.2. Conservative betting: The conservative betting policy will not allow foolish series of bets like those above. But it rules them out at the cost of disallowing bets that are obviously wise. Just switch the case around. On each toss you are offered a bet at 1:2 on $heads_i$ once you see the coin land 'p_i' or '$\sim p_i$'. Since your credence in $heads_i$ is mushy at this point you turn down all such bets. Meanwhile Sarah is looking on but makes a point of covering her eyes when the coin is tossed. Since she doesn't learn whether the coin landed

'p_i' her credence in *heads*$_i$ remains sharply ½ and so takes every bet. (Of course she does. At 1:2 the odds are strongly in her favor). Sure enough, she makes a killing. "Don't you want to get in on this?" she asks. "I can't" you reply. "I keep seeing how the coin lands, so none of these bets are rational for me." Eventually you cave and join Sarah in making money hand over fist. But of course you don't bother to close your eyes. It is not as though Sarah's lack of knowledge concerning the 'p_i' and '$\sim p_i$' outcomes has anything to do with her success. That coin is going to land *heads* about half the time regardless if anyone is watching.

As before there is a further problem for Conservative betting related Reflection. Prior to seeing the coin land you are offered a bet at 1:2 on *heads*. You are also offered the option of calling off the bet once you see it land 'p' or '$\sim p$'. According to the dilation account and Conservative betting strategy, you know that once you see the coin land you will want to renege on the bet. For you know that your credence in *heads* will dilate and the Conservative betting cautions against bets within your mushy range of credence. Indeed if you accept these you should have a policy of canceling the bet once you see the coin land. But to have such a policy and know that you will cancel the bet really amounts to rejecting the bet even before you see the coin land. But surely you don't want to do that! If your credence in *heads* prior to the toss is sharply ½—as of course it should be—then 1:2 odds are strongly in your favor. Taking such a bet ought to be a no-brainer.

Objection 4. Many Coins again: The objections involving repeated tossing from the previous section depend on particular decision theories for mushy credence. Perhaps there are other options for betting that avoid these worries.[11] But I think the basic worries can be brought out independently of decision theory. Suppose we repeat the experiment as above using a series of propositions $p_1, p_2, \ldots,$ such that $C(p_i) = [0,1]$. Upon seeing the coin land 'p_i' or '$\sim p_i$' for each toss we are supposed to have mushy credence in *heads*$_i$ for that toss. Suppose that on each toss we also get to remove the labels and see if the coin did land *heads*. It will be hard not to notice after a while that about half the time when you are in the mushy state concerning *heads*$_i$ the coin does land heads. And it will be hard not to infer *inductively* that it will continue to be the case that about half of the occasions in this scenario in which

[11] If the answer is that one should bet *as if* one has sharp credence we might wonder once again what the difference between sharp and mushy credence really consists in.

you've seen the coin land and your credence is $C(heads_i)$ = [0,1], $heads_i$ is true. Now if you know that about half of the time when you are in a certain evidential situation it turns out to be raining, then surely when you next find yourself in such an evidential situation your credence that it is raining should be very close to ½. So similarly in the case as described your credence in $heads_i$ should be ½ when you see the coin land 'p_i' or '$\sim p_i$'. But you didn't really need to learn this inductively did you? Wasn't it obvious from the beginning that this fair coin would land *heads* about half the time regardless of whether you have seen it land?

Here is a similar puzzle. I label and toss 1 million fair coins with propositions as before. This time suppose that your credence in the proposition is only *slightly* mushy, say $C(p_i)$ = [0.4,0.6]. And suppose that the p_i are *independent* across all the functions in my representor, i.e. $\forall P \in R, P(\&_i p_i) = \prod P(p_i)$. (Perhaps p_i = 'a black ball will be selected from Urn$_i$' where each urn contains 10 marbles, between 4-6 of which are black, and they have each been filled in manner causally independent of the others).

Prior to reading how any of the coins landed, what is your credence in the following proposition?

half-heads: About half the coins—say 45% – 55% of them—landed *heads*.

You know it is overwhelmingly likely that about half of a group of a million fair coins will land *heads*, so your credence in *half-heads* is very high. What about your credence in the following?

half-true: About half of the propositions $p_1, p_2, \ldots, p_{1000000}$—say 45% – 55% of them—are true.

You have a function *P'* in your representor such that for each i $P'(p_i) = 0.6$. And another $P''(p_i)$ such that for each i $P''(p_i) = 0.5$. (And there will be plenty of functions in between, assigning different values to different p_i). Now,

P'(*half-true*) must be very low.
P''(*half-true*) must be very high.

So *C*(*half-true*) = [*x*, *y*], where $x \approx 0$ and $y \approx 1$.

Now according to the dilation account, when you read how each coin landed, for all i, $C_+(heads_i) = C(p_i)$. And so C_+(*half-heads*) = *C*(*half-true*) = [*x*, *y*], where $x \approx 0$ and $y \approx 1$. That is, your credence that half the coins landed *heads* should now dilate to a highly agnostic state covering almost the whole range from 0 to 1. But now ask yourself seriously if this is what you really think. You are looking at pile of a million *evenly weighted coins, each of which has a track record of landing heads about as often as tails*, labeled with a bunch of propositions (p_1, $\sim p_2$ $\sim p_3$, p_{-4}…) about which you have no more evidence one way or another. Do you really doubt that about half of them are *heads*?

We can drive this worry further by supposing instead that for all *i*, $C(p_i) = [0,1]$. Now by reasoning along the lines above, the dilation account entails that C_+(*all the coins landed heads*) = [0,1]. That is, upon reading for each coin whether it landed 'p_i' or '$\sim p_i$' you will become maximally agnostic concerning whether *every one of a million coins landed heads*. I doubt that you are. I'll bet you're about as confident as you are about anything that they didn't all land *heads*.

5. Conclusion

I'm not happy with the mushy credence response to the evidentialist argument for POI. Either it needs to be worked out in some new subtle way, or we owe Laplace an apology for deriding his principle.[12]

References

Elga, A. 2000. 'Self-locating Belief and the Sleeping Beauty Problem' *Analysis* 60: 143-147.
Elga, A. (unpublished ms.) 'Subjective Probabilities Should be Sharp'.
Gillies, D. 2000. *Philosophical Theories of Probability*. London: Routledge Press.
Hacking 1965. *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.

---

[12] Thanks to audiences at University of Vermont, UNC Chapel Hill, Princeton, ANU, and MIT for feedback. In particular I've benefited greatly from conversation or correspondence with Rachel Briggs, David Christensen, Adam Elga, Alan Hajék, Caspar Hare, Matt Kotzen, Sarah Moss, Agustin Rayo, Susanna Rinard, Scott Sturgeon, Elliott Sober, and Bas van Fraassen.

Hacking 2000. *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Hajék, A. 2007. 'Interpretations of Probability' *Stanford Encyclopedia of Philosophy*.

Kyburg, H. 1977. 'Randomness and the Right Reference Class' *Journal of Philosophy* 74: 501-21.

Levi, I. 1974. 'On Indeterminate Probabilities' *Journal of Philosophy* 71: 391-418.

Levi, I. 1977. 'Direct Inference' *Journal of Philosophy* 74: 5-29.

Lewis, D. 1980. 'A Subjectivist's Guide to Objectice Chance' in R. Jeffrey ed. *Studies in Inductive Logic and Probability*. Vol II. Berkeley: University of California Press.

Joyce, J. 2005. 'How Probabilities Reflect Evidence' *Philosophical Perspectives* 19: 153-178.

North, J. (unpublished ms.) 'Symmetry and Probability'.

Seidenfeld, T. & Wasserman, L. 1993. 'Dilation for Sets of Probabilities' *The Annals of Statistics* 21: 1139-1154.

Strevens, M. 1998. 'Inferring Probabilities from Symmetries' *Noûs* 32: 231-246.

Sturgeon, S. 2008. 'Reason and the Grain of Belief' *Noûs* 42: 139-165.

Talbot, W. 1991. 'Two Principles of Bayesian Epistemology' *Philosophical Studies* 62: 135-150.

van Fraassen, B. 1984. 'Belief and the Will' *Journal of Philosophy* 81: 235-256.

van Fraassen, B. 1989. *Laws and Symmetry*. Oxford: Clarendon Press.

van Fraassen, B. 1990. 'Figures in a Probability Landscape', in J. Dunn and A. Gupta eds. *Truth or Consequences*. Dordrecht: Kluwer.

van Fraassen, B. 2005. 'Conditionalizing on Violated Bell's Inequalities' *Analysis* 65: 27-32.

van Fraassen, B. 2006. 'Vague Expectation Loss' *Philosophical Studies* 127: 483-491.

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

White, R. 2005. 'Epistemic Permissiveness' *Philosophical Perspectives 19*.