On Probabilistic Knowledge

John MacFarlane

May 1, 2019

Abstract

This is a lightly revised version of my comments from the Author Meets Critics symposium on Sarah Moss's *Probabilistic Knowledge* at the Central Division APA, Denver, February 20, 2019.

1 Introduction

Sarah Moss has written a terrific book. Her topics are of broad general interest, her claims novel and important, her writing clear and elegant. She weaves together themes from semantics, epistemology, and practical philosophy in ways that are genuinely illuminating. It's a book that philosophers working in many different areas will enjoy and profit from.

One way of stating the central thesis of Moss's book—the way suggested by the title—is that we can have probabilistic knowledge. We can know, for example, that it is between 50% and 90% likely to snow tomorrow, or that the defendant is more likely than not to have committed fraud, or that it is very unlikely that the lottery ticket in my pocket is the winning one.

Moss makes a strong case for the claim that making sense of probabilistic knowledge is important for a wide range of philosophical purposes. For example, if the rodeo organizers know that 60% of the people at a rodeo climbed over the fence to avoid paying the entrance fee, why can't they sue an arbitrary member of the audience—say, Elwood? After all, the standard for conviction in civil cases is "preponderance of evidence," and the statistics justify a credence of .6 that Elwood jumped the fence. Moss's suggestion is that although the credence of .6 is justified, we don't *know* that Elwood is .6 likely to have jumped the fence, because we can't rule out the possibility that he is an exception to the statistical generalization about the crowd. Talk of justified credence, then, is not enough; we need to be able to talk of probabilistic knowledge.

But why, you might ask, is the thesis that we can have probabilistic knowledge even remotely controversial? Who would deny, for example, that we can know that it is likely to snow this week? And on what grounds? Few would recognize it as a dogma of contemporary philosophy that there cannot be knowledge of matters probabilistic.

What makes probabilistic knowledge problematic—something worth writing a book to vindicate—is Moss's particular view about what probabilistic *belief* consists in. On Moss's view, to believe that it is likely to snow this week is not

- to believe that the objective probability that it will snow this week is high, or
- to believe that the frequency of snow in some relevant reference class is high, or
- to believe that the evidential probability that it will snow this week, conditional on some specific body of evidence, is high, or
- to believe that some specific person or group has a high credence that it will snow this week.

Rather, to believe that it is likely to snow is simply to have a high credence that it will snow. To put it another way: there's nothing one needs to do, besides having a high credence in snow, to count as believing that it is likely to snow. And in general, to have a probabilistic belief with a certain content is to have credences that meet a certain condition.

In this setting, the idea that we can have probabilistic knowledge amounts to the idea that states of intermediate credence, or partial belief, can constitute knowledge. And this *is* a radical idea. Historically, knowledge has been limited to the province of full belief. Those who deal in partial belief—the Bayesian tradition in epistemology—have typically talked of justification rather than knowledge. This, Moss thinks, is a mistake: the concept of knowledge has just as much applicability to states of partial credence as it does to full beliefs.

The idea that a state of partial belief can constitute knowledge is attractive for many reasons. As Moss shows, the distinctions we make between knowledge and justification in the realm of full belief have counterparts in the realm of partial belief as well. For example, one can construct analogues of Gettier cases involving partial beliefs. In addition, if we think about the *uses* we make of the distinction between knowledge and justification, we can see that a similar distinction among partial beliefs would also be useful. (I've mentioned the fence jumpers at the rodeo; the book is full of other applications.) Finally, extending the concept of knowledge into the realm of partial belief helps bring together the two hitherto largely separate worlds of traditional and Bayesian epistemology, and holds out hope that the concept of knowledge might still play a central role in epistemology even in our age of "big data."

However, there are some significant obstacles to making sense of probabilistic knowledge, if we think of probabilistic belief the way Moss does. One class of obstacles comes from the *factivity* of knowledge. The other comes from the need for a substantive "anti-luck" condition such as *safety*. I want to discuss each of these problems and the way Moss deals with them. But before doing that, I'll need to say a bit about Moss's elegant and innovative approach to the semantics of probabilistic language and the content of probabilistic thought.

2 Semantics

2.1 The contents

On Moss's view, probabilistic contents don't have possible-worlds intensions. They don't classify classify possible states of the world; instead, they classify *information states*. Moss models information states as probability spaces. The content of a probabilistic belief, then, is a set of probability spaces. The content *it is likely to snow*, for example, is the set of probability spaces that assign a high probability to snow.

I am very sympathetic to this idea, having defended a view myself on which contents are information-sensitive (MacFarlane 2011, 2014, ch. 10; Kolodny and MacFarlane 2010; see also Yalcin 2007, 2011). But there's a significant difference between Moss's approach and mine. While Moss models contents as sets of information states, I model them as sets of world/information state *pairs*. The motivation for using pairs is to make sense of contents that are partly probabilistic and partly worldly, such as

(1) Either the snow plows went out or it's likely that there were many accidents.

On my view, the truth of the second disjunct varies with the information state, while the truth of the first disjunct varies with the world; accordingly, the truth of the whole disjunction varies with both factors. Moss has a different approach. She introduces an implicit C operator (for *certainty*) that type-shifts regular contents to informational ones. On her view, then, (1) is tantamount to:

(2) Either *it's certain that* the snow plows went out or it's likely that there were many accidents.

Both disjuncts are now sensitive only to the information state, not to the world, and the content of the disjunction can be a set of information states.

2.2 Partitions

You might have the following worry: in a typical context where one would assert (1), one won't be certain that the snow plows went out. But in such a case the first disjunct of (2) would be known to be false. In most contexts it is infelicitous to assert a disjunction one of whose disjuncts is known to be false. How, then, can (1) can be felicitously asserted, on Moss's analysis?

To explain this, I need to introduce an additional feature of Moss's semantics. On Moss's view, every epistemic modal, probability operator, conditional, and sentential connective has an index that gets filled in context with a set of propositions that partitions logical space. Following Hamblin (1973), we can think of such a partition as the set of possible answers to a *question*.

The content of a disjunction *p or q* is the set of probability spaces such that the result of updating them with *any* of the propositions in the partition would belong to one of the disjuncts. Intuitively: one believes *p or q*, relative to a partition X, if one would believe either p or q (or both) after learning *any* of the propositions in X.

If the partition is just the trivial partition—a single proposition encompassing all of logical space—then (2) is equivalent to a truth-functional disjunction, and is not felicitously assertable in a context where the first disjunct is known to be false. But suppose the partition attaching to *or* is

PlowsOut? {the snow plows went out, the snow plows didn't go out}.

Then (2) is not equivalent to a truth-functional disjunction. We believe (2) (relative to this partition) because, if we learned that the snow plows went out, we'd accept its first disjunct, and if we learned that they didn't go out, we'd accept its second disjunct. So understood, with the disjunction relativized to *PlowsOut?*, (2) is felicitously assertable even when its first disjunct is presupposed to be false.

Moss also uses this partition-relativity to explain how we can get non-collapsed readings of iterated epistemic modal and probability operators. Consider

(3) It might be probable that Kenny is here.

According to the semantics for epistemic modals I have advocated, this is just equivalent to

(4) It is probable that Kenny is here.

For (4) imposes a condition on information states alone: it does not vary in truth from one world to another, keeping fixed the information state. It is true at (w, i), then, just in case it is true at (w', i) for any w'. If (4) might be the case, then, it *is* the case, so (3) is not weaker than (4).

You might think that this is just a bad prediction of my theory. Isn't it obvious that (3) *is* weaker than (4)? In *Assessment Sensitivity* (MacFarlane 2014, sec. 11.7), I give an error theory to explain away this judgment. I suggest that we judge (3) to be weaker than (4) because we think we can envision scenarios where (3) is true but (4) is not. In these scenarios, we're ignorant of some plain matter of fact on which the truth of (4) hangs. For example:

(5) If Kenny finished his book, he is probably here.
(6) If Kenny didn't finish his book, he probably isn't here.
(7) Kenny might have finished his book, and he might not have.

We think it follows that

(8) It might be that Kenny is probably here, and it might be that he probably isn't here.

which would support (3) but not (4). However, this sort of argument is invalid on the semantics I give for conditionals and epistemic modals (see MacFarlane 2014, ch. 10; Kolodny and MacFarlane 2010). Because it shares a form with arguments that are clearly good, people are naturally led to the mistaken conclusion that (8) is true and that they are in a position to believe (3) but not (4). These judgments are errors—but understandable ones.

On Moss's view, by contrast, there is a reading of (5–8) on which it is a good argument, so these judgments cannot be dismissed as errors. To get this reading, one must index *probable* and *and* with the trivial partition, and *might* with the partition

Finished? {Kenny finished his book, Kenny didn't finish his book}.

Then we accept *It might be that Kenny is probably here* because there is a proposition in *Finished?* such that if we learned it, we'd accept that Kenny is probably here. And we accept *It might be that Kenny probably isn't here* because there is a proposition in *Finished?* such that if we learned it, we'd accept that Kenny probably isn't here. On this reading of the modals, (3) is indeed weaker than (4), and the argument (5–8) supports (3) but not (4).

This is really an elegant approach, and there may be reasons to favor it over my error theory (which Moss doesn't discuss explicitly). But I do want to signal two reservations I have about it. The first is that it overgenerates readings. The second is that it posits indices for which there is no independent motivation.

2.3 Overgeneration

A fair six-sided die has been rolled; we haven't seen the outcome. So, presumably we should believe

(9) If the die landed on a low number (1–3), it probably landed on an odd number.

On Moss's view, though, both *if* and *probably* are indexed with a contextually supplied partition. If we interpret the conditional in (9) relative to the partition

> *LowOrHigh?* {low, high},

then the content of (9) includes our current credal state (that is, we believe it). But if we interpret the conditional relative to the partition

> *WhichFace?* {1,2,3,4,5,6},

then the content of (9) does not include our current credal state, because there is a proposition in this partition, namely 2, that accepts the antecedent but not the consequent.

Both *LowOrHigh?* and *WhichFace?* are natural partitions to think about in this case.[1] So if Moss's view is right, we should expect to be able to hear both readings. But I find it difficult to hear the reading with *WhichFace?*, even with ample contextual prompting. Consider this dialogue:

> Crito: If it's low, it's probably odd.
> Euthydemus: You don't know that! After all, one way it could be low is
> by being 2, and in that case it would definitely be even.

Euthydemus just sounds like a sophist here, and his intervention doesn't make me any more inclined to reject the conditional.

Similarly, if the partition *WhichFace?* is available, it should be okay to say

(10) It might be probable that the die came up 4,

[1] In addition, both satisfy the presupposition of being *decisive* with respect to the antecedent (Moss 2016, 73): each element of the partition entails either the antecedent or its negation.

without implying that the die might be loaded. I want to invite Moss to explain why this reading isn't readily available here—even if we try to raise *WhichFace?* to salience by saying, "the question is which of its six sides the die landed on."

2.4 No independent motivation for the indices

I think we should be suspicious when semanticists posit variables that can *only* be filled by context, and not given values explicitly or bound by quantifiers. On Moss's view, every modal, conditional, and sentential connective has an index that is assigned a partition as its value in context. But we don't seem to have any natural way to say "it is likely relative to such-and-such a partition that p", or "if (relative to such-and-such a partition) p, then q." I can't make the true reading of (10) any more salient by saying

(11) Speaking to the question which of its six sides the die landed on, it might be probable that it landed on 4.

And I can't dispel the contradictory impression made by

(12) It's not probable that the die landed on an even number, but it might be probable that the die landed on an even number

by saying

(13) Relative to the question how I should bet, it's not probable that the die landed on an even number, but relative to the question whether the die landed low or high, it might be probable that the die landed on an even number.

If these constructions are really partition-sensitive, why don't we have ways of making their implicit completions explicit? This definitely sets these indices apart from quantifier domains, modal bases for epistemic modals, indices on pronouns, etc.

3 Epistemology

Let's now turn from semantics to epistemology.

3.1 Factivity

Knowledge is factive. Does that pose a problem for the idea of probabilistic knowledge?

Formally, factivity amounts to the entailment

(14) S knows that $p \rightarrow p$.

Given (14), believing that someone knows a probabilistic content requires believing that content yourself. It follows that attributions of probabilistic knowledge will have probabilistic contents too. But that is a consequence Moss is willing to accept.

Moss explores some ways in which factivity might be used to mount a skeptical argument against probabilistic knowledge. Here is one. Imagine that I have confirmed the fairness of a coin in every way possible, then flipped it, but not looked at the result. I believe

(15) The coin is .5 likely to have come up heads.

Intuitively, this should be something I know, if I have any probabilistic knowledge at all. But although I haven't looked at the coin, I know that

(16) Either it landed heads or it landed tails.

I also know that if it landed tails, there is no chance it landed heads. Thus,

(17) If it landed tails, then it is not .5 likely it landed heads.

By the same token, if it landed heads, it is not .5 likely to have landed tails. But if it's not .5 likely to have landed tails, then it's not .5 likely to have landed heads. So, I know that

(18) If it landed heads, then it is not .5 likely it landed heads.

Putting (16), (17), and (18) together, I might conclude that

(19) It is not .5 likely it landed heads.

And since knowledge is factive, it seems to follow that I can't know (15) after all. Generalizing this reasoning, one could argue that only non-probabilistic contents can be known.

Moss argues that this argument is unsound (cf. Moss 2016, 142). In its intended interpretation, (16) is

(20) C(it landed heads or it landed tails)

but what is needed for the argument to be valid is

(21) C(it landed heads) or C(it landed tails).

I think this is the right approach (the argument is like the miners case discussed in Kolodny and MacFarlane 2010). But there remains a difficulty that this kind of logical maneuvering can't paper over. Suppose I know that it is .5 likely that the coin landed heads. It seems compelling to say that, when I look at how the coin came up, I'll come to believe either that it is certain that the coin landed heads, or that it is certain that it landed tails, and that my belief will constitute knowledge. Either way, I'll come to know that it's not .5 likely that it landed heads.

But given factivity, we can't coherently say both that I know now that it is .5 likely that the coin landed heads, and that after examining the coin I'll come to know either that it certainly landed heads or that it certainly landed tails. If I take myself to know now that it's .5 likely that the coin landed heads, then given factivity, I have to deny that I'll come to *know* otherwise when I look at the way the coin landed. But I can't deny that I'll come to *believe* otherwise. So, I'll have to suppose that I'm going to go off the rails, when I examine the coin, and come to believe something that isn't true. (Once I've looked at the coin, of course, I'll switch perspectives: I'll deny that I had knowledge before.)

Of course, the same kinds of worries arise for the sorts of relativist views I have defended (in MacFarlane 2014 and elsewhere). But I think it helps to have the explicit relatization of truth to contexts of assessment, which allows us to formulate norms for assertion and belief that make sense of my future self's epistemic behavior (MacFarlane 2014, ch. 5, §12.1). Although he will be forming a belief that is false relative to my perspective, it will be true relative to his, so what he is doing is, in an intelligible sense, objectively correct. (He is conforming to the objective norm for believing.)

However, Moss rejects any explicit relativization:

> my probabilistic theory of epistemic vocabulary does not make use of any relation of relative truth. The criminal *may be justified* in believing the content that Bond is probably in London. But the content of her belief is not true for her, or indeed true in any sense. The content is inconsistent with the true content that Bond is not in London, and that alone is enough to guarantee that the content is false. (Moss 2016, 128, emphasis added)

The thought here is that we can appeal to notions like justification to explain how the probabilistic beliefs formed by others with different information are in some sense right. But it's hard to see how, without explicit relativization, I can even sustain the idea that my future self will be *justified* in holding that it is not .5 likely that the coin landed heads. To say that Future Self is justified is to say that he forms his belief in a way that could be expected to lead to true beliefs. But how can I see him as doing that? After all, he forms probabilistic beliefs that conform to *his* information, but (by my lights) the truth of these beliefs depends on *my* information. So, his process of belief formation is, by my lights, completely unreliable (cf. MacFarlane 2014, sec. 7.2.8). Even when he does manage to get things right, by my lights, it's hard to see this as anything but a lucky accident. (And what goes for Future Self goes for other people too.)

3.2 Safety

Knowledge, many people think, requires safety: if you know that p, then you couldn't have easily been wrong. Typically this is spelled out in terms of possible worlds:

(22) S's belief that p is *safe* iff there is no close world where S believes p but p is false.

But how can we make sense of this for probabilistic contents, which classify probability spaces rather than worlds? What sense can there be in asking whether a probabilistic content would be true if the world were a different way? The idea of truth at a world just doesn't seem to make sense for contents that are true relative to information states, not worlds.[2]

Moss's approach to this problem is subtle and resourceful:

> As it turns out, there is a sense in which *probabilistic contents* can be true or false at possible worlds. For instance, the content that it is .6 likely that Jones smokes can be true at a world—namely, just in case that world is .6 likely to be a world where Jones smokes. This particular truth condition is just another probabilistic content, namely the content that you believe when you have .6 credence that Jones smokes at the world in question. (p. 155)

The idea is that a "truth at a world" claim like

(23) the proposition (that it is .5 likely that p) is true at w

is equivalent to a probabilistic content

[2] The problem does not arise in the same way if we model probabilistic contents as sets of world/information state pairs, for then it makes sense to evaluate the content at different worlds, keeping the information state the same. On such a view safety is *trivially* satisfied for contents whose truth depends only on the information state component of the pair. But is this a problem? Even those who promote safety take it to be only a necessary condition for knowledge, not a sufficient one. In any case, probabilistic beliefs would be in good company: true mathematical beliefs trivially satisfy safety, as formulated in (22), and nobody thinks this poses a problem for the idea of mathematical knowledge.

(24) w is .5 likely to be a world where p is true.

Here p is an ordinary, non-probabilistic proposition, so we have eliminated talk of truth-at-a-world for probabilistic contents in favor of a probability operator and truth-at-a-world for an ordinary content.

Although Moss doesn't give a precise algorithm for performing these reductions for arbitrary propositional contents, it does seem plausible that there is a way to do this. For example,

(25) the proposition (that p or probably q) is true at w

would presumably reduce to

(26) either w is a world where it is true that p or w is probably a world where it is true that q.

Presumably, we also need to preserve the partitions assigned to the indices on each operator. (If not, it needs to be clearly stated how they change.) But I wasn't sure this could be done. For in the transformation from (25) to (26), we change the subject matter of the content to which the probability operators and sentential connectives apply. For example, if (as in Moss's own example) p is *the president will be Trump* and q is *the president will be Clinton*, then the partition attaching to *or* might be something like

> *WhichParty?* {The Republican candidate will win the general election, the Democratic candidate will win the general election}

But *WhichParty?* can't be the partition attaching to *or* in (26), which is a proposition *about a world*. The partition there should be something like

> *WhichPartyInw?* {The Republican candidate will win the general election in w, the Democratic candidate will win the general election in w}

and this is not the same partition as *WhichParty?*. I may be missing something here, so I'll just see if Moss wants to say anything about this issue.

Having shown how to formulate safety for probabilistic beliefs, Moss considers a skeptical argument that even paradigm probabilistic beliefs will not satisfy safety. Consider our flipped coin again. I believe that (15) it is .5 likely that it landed heads. But, however it landed, there's presumably a close possible world w_T where it landed tails. Is (15) true at w_T? The skeptic asks: how could it be? As we have already seen, to say that (15) is true at w_T is to say that

(27) w_T is .5 likely to be a world where the coin landed heads.

But by stipluation, w_T is a world where the coin landed tails. So, w_T is *definitely* not a world where the coin landed heads. And this means that (27) is false. It is false, then, that (15) is true at w_T, so my belief in (15) is not safe.

In response, Moss observes that probabilistic claims about objects—including possible worlds—are hyperintensional (Moss 2016, 149). Whether this particular lottery ticket is unlikely to have won depends on how the ticket is described. Suppose that the ticket is, in fact, the winning one (though we don't know this), and that it was dubbed "Winner" on TV when it was selected last night. Under the description "Winner," it is not at all unlikely that this ticket won. Under the description "Ticket #23433", on the other hand, it may be very unlikely that it won.

Appreciating this hyperintensionality is the key to responding to the skeptic, Moss thinks. For whether (27) holds depends on how we think of or describe w_T. If we think of w_T as *the world where everything is as much as possible like actuality, but the coin landed tails*, then (27) is false and (15) is false at w_T. But if we think of w_T as *the world where everything is as much as possible like actuality, but the coin was flipped a bit harder*, then (27) is true and (15) is true at w_T.

The upshot is that whether my belief that (15) is safe—hence, whether it amounts to knowledge—depends on how we conceptualize close possible worlds like w_T.[3]

This response seems to me to concede too much to the skeptic, by granting that focusing a certain description of a nearby possibility can make it impossible to attribute knowledge that a lottery ticket is probably a losing one. Consider the absurdity of this dialogue:

> CRITO: I probably lost.
> EUTHYDEMUS: You don't know that! Whether or not you lost in actuality, your ticket could easily have been been the winner (just as easily as any other). And in that case it wouldn't have been true that you probably lost. Since you could easily have been wrong, you don't have knowledge.

It would be entirely natural for Crito to reply, mystified:

[3]Note that as formulated in (22), safety is a quantified claim. On Moss's view, the hyperintensionality of probability operators extends to the quantifiers. The truth of a quantified sentence like *Every ticket (is such that it) probably lost the lottery* (Moss 2016, 151) can depend on how we conceptualize the domain: if the domain is conceived under modes of presentation *ticket 1*, *ticket 2*, and so on, it is true; but if the domain is conceived under modes of presentation *winning ticket*, *losing ticket 1*, and so on, it is false. Moss suggests that the assignment function would assign *senses* rather than objects to the bound pronouns, but she gives no details about how the hyperintensionality is to be implemented in her semantics.

> CRITO: It was because I imagined this possibility so vividly that I bought the ticket in the first place. Nonetheless, I knew, and still know, that the ticket probably isn't the winner.

Notice, though, that the skeptical argument Moss is trying to meet with the invocation of hyperintensionality is a creature of her own gloss on safety. She could avoid it entirely if she just conceded that safety—which is defined in terms of truth at a world—doesn't make sense as applied to contents that classify information states, not worlds. This would leave her with the task of formulating a different anti-luck condition for probabilistic knowledge. But it would avoid shoe-horning this task into the quixotic one of making sense of safety, and truth at a world, for probabilistic contents.

References

Hamblin, C. L. 1973. "Questions in Montague English." *Foundations of Language* 10: 41–53.

Kolodny, Niko, and John MacFarlane. 2010. "Ifs and Oughts." *Journal of Philosophy* 107: 115–43.

MacFarlane, John. 2011. "Epistemic Modals Are Assessment Sensitive." In *Epistemic Modals*, edited by Andy Egan and Brian Weatherson. Oxford: Oxford University Press.

———. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.

Moss, Sarah. 2016. *Probabilistic Knowledge*. Oxford: Oxford University Press.

Yalcin, Seth. 2007. "Epistemic Modals." *Mind* 116: 983–1026.

———. 2011. "Nonfactualism About Epistemic Modality." In *Epistemic Modals*, edited by Andy Egan and Brian Weatherson, 265–332. Oxford: Oxford University Press.