Economics & Philosophy (2024), **40**, 507–511
doi:10.1017/S0266267124000051

CAMBRIDGE UNIVERSITY PRESS

SYMPOSIUM ARTICLE

Replies to commentators

Anna Mahtani

Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK
Email: a.mahtani@lse.ac.uk

(Received 31 December 2023; accepted 3 January 2024; first published online 27 March 2024)

I'm so grateful to the commentators for their insightful and constructive responses! Below I continue this exchange with a brief note of reply.

Richard Bradley focuses on the Principal Principle, discussed in section 5.3 of *Objects of Credence*. Bradley begins by arguing that cases of the contingent a priori do not work as counterexamples to the Principal Principle. The underlying thought is that there are 'different possible ways of mapping language to a propositional framework', and I see this thought as congenial to 2-D semantics, on which the primary intension of *Lucky wins* is the set of all possible worlds, while the secondary intension of *Lucky wins* is the set of those worlds where, say, Alice wins, if it turns out that Alice and Lucky are one and the same. The thought then is that once the set of worlds is specified precisely, the chance and credence functions will assign it the same credence, and the Principal Principle will be vindicated. We might add that the apparent problem for the Principal Principle arose because we generally interpret a chance claim as assigning a value to the secondary intension of a proposition, and a credence claim as assigning a value to its primary intension. I think this is a promising and interesting idea, and my main objection to it is that (as I argue in Chapter 7 of *OOC*) we face some challenging complications if we take the objects of credence to be either primary or secondary intensions.

Bradley turns then to my 'New New Principle', according to which we should defer to the chance function conditionalized on the a priori. Bradley goes further and suggests that the chance function ought in any case to be conditionalized on the a priori – in which case no amendment to the Principal Principle is needed. But I argue that the chance function conditionalized on the a priori is omniscient: that is, it assigns 1 to every true claim, including claims about the currently unsettled future. Here's one way of running the argument:

(1) If P is true, then *actually* P is necessary.
(2) If a claim is necessary, then chance assigns it a value of 1.
Sub-conclusion: Therefore, if P is true, then chance assigns 1 to *actually* P.

(3) It is a priori that *P iff actually P*.[1]
Conclusion: If *P* is true, then chance conditionalized on the a priori assigns a value of 1 to *P*.

Bradley disagrees with the conclusion of the argument. Which premise does he reject? It is clear that the sub-conclusion is rejected, for Bradley writes 'if the chance of *P* is *x*, then the chance that *P* is actually true is also *x*'. This is incompatible with the sub-conclusion, assuming that there is some claim *P* that is true, but assigned a chance less than 1. Bradley must then reject either (1) or (2).

To reject (1) would be to reject a standard account of the 'actually' operator, on which if *P* is true at the actual world, then *actually P* is true at every possible world. To reject (1) would (again) resonate with 2-D semantics, according to which the primary intension of *actually P* is the set of worlds at which *P* is true. But it seems that Bradley's strategy is rather to reject (2), for he writes: 'it does not follow from the fact that *actually P* is either necessarily true or necessarily false that its chance should be either zero or one'.[2] Bradley's strategy then, as I understand it, involves the interesting and radical move of rejecting the Basic Chance Principle, which connects chance and possibility.

Luc Bovens' response focuses first on the reflection principle (section 5.2 of *OOC*). I suggest that the reflection principle should require deference (over some proposition P) to an agent (under a designator) only if the designator is 'appropriate' – that is, only if were the agent to learn that they are so-designated, their credence in P would be unaffected. This allows us to explain (amongst other things) the failure of the standard reflection principle in the card game case (section 5.2.2) and in the Sleeping Beauty problem (section 5.2.4). Bovens argues that these admit of simpler solutions.

For the card game case, Bovens suggests that we should adopt his 'Clarified Generalised Reflection Principle$^{\#'}$', which states that if a rational agent-at-a-time A respects a single agent-at-a-time A* (so designated), then A defers to A* (so designated), and if A respects multiple agents-at-times (so designated), then A's credence equals the mean of these agents-at-times' credences. But consider a variation of the card game case in which you are quite sure that the mug will remain rational throughout, but have a slight doubt about both the lucky player and the other player: perhaps you think that seeing an ace can occasionally make people less rational (giddy with the expectation of a win perhaps). In this case, at t_0 you respect a single agent-at-time-t_1 – namely, the mug – and so Bovens' clarified generalised reflection principle requires you defer to that agent, and have a credence of ⅓ in two-aces. But that can't be right! Some sort of amendment would be needed to the principle to avoid this conclusion.

[1]From this premise it follows that chance conditionalized on the a priori assigns the same value to *P* as it assigns to *actually P*.

[2]Similarly, later in his response Bradley writes: "although prior to the coin landing either necessarily Lucky is Ann or necessarily Lucky is Bob, it is not settled which it is; hence one cannot infer that all reasonable chance functions assign measure zero or one to these propositions". This entails that a claim can be necessarily true/false (at *t*) and yet not be assigned a chance of 1/0 (at *t*), which is again a rejection of the Basic Chance Principle.

https://doi.org/10.1017/S0266267124000051 Published online by Cambridge University Press

Bovens offers a different response to the Sleeping Beauty case. The thought is that SB loses information *about what day it is* between Sunday night and Monday morning, and that this loss of information explains why SB on Sunday night should not defer to herself on Monday morning. But consider a variation in which SB arrives at the laboratory without knowing whether it is Saturday or Sunday – just knowing that she will now be put to sleep and woken on Monday, at which point the experiment will proceed in the usual way. Now we have our puzzle once more: why doesn't SB when she arrives (presumably with a credence of ½ in HEADS) defer to her Monday-morning self (whose credence in HEADS – according to the thirder – is ⅓)? We can no longer say that between these two times SB loses information about what day it is – for she does not know what day it is at the start.

Perhaps Bovens might reply that even if SB does not know whether it is Saturday or Sunday, she still knows that it is day 0 'in the sequence' – for there is no important distinction between these two days. The thought here would be that some self-locating information matters (when assessing a given proposition) and some self-locating information does not. This is part of what the idea of an 'appropriate' designator is designed to capture: you need to know that you fall under a designator iff it matters for the resolution of the proposition under consideration.

Finally Bovens turns to the *ex ante* pareto principle. Bovens points out that under my supervaluationist interpretation, there are cases where we would expect the *ex ante* pareto principle to apply – and where we feel like it *should* apply – but where it does not. I agree. Does it follow that we should reject the supervaluationist interpretation? Sticking with the old *ex ante* pareto principle is (I claim) not an option, because – given the tenet of *OOC* – it is simply incoherent. Thus the principle must be re-thought, and a supervaluationist interpretation of the principle seems like the natural move. In line with Bovens' example, I see many cases where the *ex ante* pareto principle so interpreted will make fewer rulings than we might have expected.

Melissa Fusco offers a Stalnaker-style analysis of some of the problems discussed in *OOC*. To give a very rough summary, the idea is that the objects of credence are 'diagonal' propositions: the 'diagonal' proposition corresponding to an utterance is the set of possible worlds w such that the utterance – as made in w, considered as actual – is true at w.

There seems to me to be much to recommend this idea, but there are cases where I find the proposal troubling. Let's consider the Sleeping Beauty case. According to the thirder SB's credence in HEADS is ½ on Sunday night, but ⅓ on Monday morning. What proposition p did SB learn between these two times? Standard conditionalization does not seem to apply. If SB learns anything, she might express what she has learnt on Monday with the assertion 'I am awake today' – but it's hard to see how she can have conditionalized on this, for what credence did she have on Sunday night in the content of this assertion (as uttered on Monday)? Fusco argues that what SB learns is the diagonal proposition – roughly that 'I am awake today' is *true* – and that by conditionalizing on this we can explain the thirder's result. But just as it was unclear how SB on Sunday night could have a credence in the content of 'I am awake today' (as uttered on Monday), so it seems unclear how SB on Sunday night could have a credence in the relevant diagonal proposition.

https://doi.org/10.1017/S0266267124000051 Published online by Cambridge University Press

How could SB on Sunday night think about the Monday morning assertion, except as an assertion *made on Monday*?

Fusco also discusses the two-envelope paradox, and in a very interesting argument, Fusco suggests that sometimes it is the value of a *diagonal* proposition which is relevant to decision-making. Let us imagine that I have the envelope I have selected in my hand, and I'm wondering whether it would be wise to stick or switch. The value of sticking can be given by the assertion 'I get the envelope I actually have', but Fusco argues that I may be unsure of the content of that assertion. Does it mean that I get the envelope containing the lesser amount, or the envelope containing the greater amount? By focusing on the value of the diagonal proposition, Fusco derives the right decision in this scenario. But my worry here is that it seems a bit of a stretch to say that I am unsure of the content of the assertion 'I get the envelope I actually have'. Surely I know what object 'the envelope I actually have' refers to (the envelope right here in my hand!) even if I don't know its monetary worth, just as I know what it refers to even if I don't know its weight? In general there seems to be a problem with the diagonalization strategy as a response to the problems in *OOC*, for in many of the cases discussed it does not seem as though our ignorance is even partially linguistic: Tom can know to whom 'George Orwell' refers (that famous author) and also know to whom 'Eric Blair' refers (that customer in the café) even if he does not know that 'George Orwell' and 'Eric Blair' co-refer.

Christian List sets out a framework that is designed to accommodate *OOC*'s tenet – that the objects of credence are opaque – while also retaining the standard Bayesian assumption that the objects of credence are elements of some algebra. List begins by defining a language as a set of elements – which we can call sentences – endowed with a negation operator. List describes the process of *engineering* a notion of tenability – where a subset of L can be either tenable or untenable. On a natural interpretation of tenability, List writes 'any set of sentences from L whose inconsistency or mutual incompatibility an agent is not – or not currently – able to establish could be deemed tenable'.

From here, List constructs a set of worlds or states (which are particular sorts of subsets of L), and defines propositions as sets of such worlds: a sentence L will express some such proposition. With this machinery, List can define an algebra over the propositions, and a probability function on that algebra. Thus we have a framework that is very amenable to the standard Bayesian approach.

This is the sort of framework that I was aiming towards in Chapter 8. I have just two questions or observations. Firstly, I note that List's framework is designed to accommodate an agent for whom the following set of sentences would be tenable: 'the set consisting of Peano's axioms and the negation of some complicated theorem of arithmetic'. But it may be that there is no *maximal* tenable set containing both Peano's axioms and the negation of some complicated theorem, because the agent would be able to establish that any such set is inconsistent. Is this in conflict with the completability principle, which requires that any tenable set has a tenable superset containing a member of each sentence-negation pair in L? Secondly, I just note that the framework is – as List writes, abstract and flexible. It leaves open all the important decisions about its interpretation. Language L cannot, I think, be any natural language where context plays a role in fixing the content of an utterance, for here we are assuming that each sentence has a unique content. How can we translate

https://doi.org/10.1017/S0266267124000051 Published online by Cambridge University Press

between L and, say, English, to determine which credence attribution claims are true? Furthermore, the definition of tenability leaves much to be decided: if we allow an agent's whims to entirely dictate which sets are tenable, then what – if anything – does rationality require of an agent? For many of the purposes to which the credence framework is put, these details will need to be filled in, and with List's framework in place, the scene is set for users of the framework to address these difficult questions.

Anna Mahtani is a Professor in the Department of Philosophy, Logic and Scientific Method, at the London School of Economics and Political Science (LSE). She works in decision theory, formal epistemology, philosophy of language and welfare economics.

Cite this article: Mahtani, A. (2024). "Replies to commentators." *Economics and Philosophy* **40**, 507–511. https://doi.org/10.1017/S0266267124000051