The Objects of Credence, by Anna Mahtani. Oxford: Oxford University Press, 2024. Pp. xii - 224.

Consider the following two claims:

(1) Tom has a credence of 0.8 that George Orwell is a writer.

(2) Tom has a credence of 0.1 that Eric Blair is a writer.

These credence attributions seem as though they might both be true. For Tom may be unaware that Eric Blair and George Orwell are the same person. However, if the objects of credence are propositions, conceived of as ordinary sets of possible worlds, Tom cannot rationally maintain this set of attitudes. These sorts of cases suggest that in assessing an agent's credences, it matters how the objects of credence are designated. It matters that Tom's attitudes aren't towards George Orwell or Eric Blair, but towards this individual under certain modes of presentation.

This is the central claim advanced by Anna Mahtani in *The Objects of Credence*. The first part of the book defends the claim that credence attributions are hyperintensional or 'opaque', in the way brought out by the previous example. Mahtani calls this claim 'the tenet'. The second part of the book explores the implications of the tenet for principles of theoretical and practical rationality. The third and final part of the book considers the difficulties that arise when we attempt to modify the foundations of the credence framework so as to accommodate the tenet. The book ends with the suggestion that we continue to use the credence framework, though with a heightened awareness of the complications that this use involves.

Mahtani is deliberate about structuring her discussion in a way that engages with both philosophers of language and users of the credence framework, whom she takes to include social scientists, statisticians, and policy makers. The book is meticulously argued and its conclusions are compelling. I strongly recommend it. In what follows, I'll review some of Mahtani's arguments and raise a few concerns, focusing especially on the rational constraints whose failures are supposed to motivate the tenet. One theme of my criticisms is that some of Mahtani's innovations address problems that can be accommodated by principles we already have. A second idea that I return to throughout my discussion is that certain users of the credence framework would likely take issue with the way Mahtani chooses to draw its boundaries.

1 Credence Claims are Opaque

After an overview of the debate about the nature of propositions in chapter two, and the credence framework in chapter three, the argument of the book begins in chapter four where Mahtani defends the claim that credence attribution statements create an opaque, or hyperintensional, context. Our choice behavior, including our betting behavior, supports this tenet. I might be willing to accept a bet with unfavorable odds that will pay out if George Orwell is a writer, but require much more favorable odds for a bet involving the proposition that Eric Blair is a writer. In addition to this, denying the tenet commits one to a form of factual omniscience that is at odds with the central aim of the credence

framework, which is to model uncertainty. Denying the tenet also yields counterintuitive results if we think that an agent should update by conditionalization.

In the second part of chapter four, Mahtani considers whether accounts of propositions that support the tenet can be generalized to credence claims. She begins with Salmon's view, which in introducing a three-place *BEL* relation that holds between a person, a proposition, and a guise—where a guise is 'something like a Fregean sense'— allows us to maintain a difference in our attitude ascriptions involving the same individual. (p.62) Mahtani considers how one might extend this idea to a four-place *CRED* relation that holds between a person, a proposition, a number, and a guise, and which can capture that Tom's credence that George Orwell is a writer is high under the 'George Orwell is a writer' guise, but low under the 'Eric Blair is a writer' guise. (p.66)

Mahtani notes that several of the accounts that result from attempting to extend Salmon's approach in the previous way conflict with probabilism and conditionalization. Putting these considerations to the side, I want to take a step back and begin with a more general observation about how Mahtani seems to understand the dialectic that these credal norms set up. Mahtani notes in the book's introduction that while there is agreement among users of the credence framework that credence claims are opaque, there is the tendency to rebuff the significance of this observation with the following response:

> It may well be that credence claims are opaque as you say, but the credence framework would be able to accommodate this fact without difficulty. Thus we can continue to use the credence framework just as we do now, while in parallel a largely unconnected project can be run on how to interpret the framework so as to accommodate the point that credence claims are opaque. (p.5)

Mahtani disagrees:

> On my view, the tenet forces us to reassess many of the uses to which the credence framework is put: there are concepts which need to be rethought, moves which turn out to be invalid, and principles which need to be rejected or transformed. We cannot then continue to use the framework without paying attention to this issue. (p.5)

It's not clear to me that Mahtani's discussion properly interprets and responds to the sense in which users of the credence framework take accommodating the tenet to be an unconnected project. One position, which I do not take to be widely held, is that one can deploy the credence framework in isolation from problems that arise in virtue of the tenet. However a different position, which I think many users of the credence framework *do* endorse, acknowledges that norms like probabilism and conditionalization dramatize the difficulties that arise in virtue of the tenet, while falling short of attributing these problems *to* these credal norms. This view does not deny that the tenet has implications when paired with the credence framework. What it denies is that these problems are endogenous to this framework.

The argument of the book seems to straddle these two targets. On the one hand, Mahtani uses problems with the Reflection Principle (which tells us to defer to our future self), and the Principal Principle (which tells us to defer to the chance function), to motivate

the need to adopt a non-standard account of the objects of credence in chapters seven and eight, suggesting that she is targeting the first position. On the other hand, she calls for revisions of the Reflection Principle and the Principal Principle in chapter five, suggesting that she takes these norms *themselves* to be the source of the problem. This second set of arguments suggests that the problem originates from within the credence framework. Again, however, many users of the credence framework would likely insist that the difficulties Mahtani highlights arise from the violation of general norms in the philosophy of language that apply just as readily to full belief, and which are thus independent of the credence framework. The crucial point is that this metanormative dispute, which seems to be the only genuine one, cannot be adjudicated by appealing to our intuitions about particular cases, which is the methodology appealed to throughout the book.

None of this is to suggest that the results of Mahtani's discussion aren't both illuminating and important. It is, however, to ask whether the problems she considers are ones that users of the credence framework would not already acknowledge, and whether there is not a still more pressing question that remains unaddressed.

2   Mahtani on the Reflection Principle

As noted already, chapter five considers the implications of the tenet for rational constraints on credences. Mahtani observes that deference claims, like credence claims, are opaque. I can rationally defer to someone under one guise while failing to defer to them under a different guise. This poses problems for the Reflection Principle. It poses problems even for the following more general and refined version of this principle:

> *Generalized Reflection Principle*: If a rational agent-at-a-time A respects an agent at-a-time A^*, then A defers to A^*.

To illustrate that deference claims—and, indeed, 'respecting' claims—are opaque, Mahtani appeals to the following situation:

> TWO ACES. This situation involves a dealer, three players (Alice, Bob, and Carol), and a pack of cards...The dealer randomly and secretly selected one player to be 'the mug' and one to be 'the lucky player', and then the remaining player is simply 'the other player'. The dealer deliberately deals a joker to the mug, an ace to the lucky player, and deals either a joker or an ace—chosen at random from a deck consisting of just aces and jokers in equal proportion—to the other player. All the players understand this set-up, but none knows how the roles (the mug, the lucky player, the other player) have been assigned. (p.77)

Given this way of setting things up, I can respect the same individual under, say, both the designator of Bob and the mug. But then a problem arises. If I respect both the mug and the lucky player at t_1 after they have looked at their cards, then I will be forced to defer at t_0 to their credences. This will mean maintaining a credence of both 1/3 and 2/3 in the proposition that *two of the cards dealt out are aces*. In addition to the fact that neither of these seem like the right values—my t_0 credence should intuitively be 1/2—I obviously cannot adopt them both.

3

According to Mahtani, the solution is to distinguish designators that are *appropriate* from designators that are not. A designator A^* is appropriate for A to defer to iff A is certain that either a) A^* recognizes with certainty that she falls under that designator or, b) if A^* were to learn just that she is A^*, then A^*'s credence in the proposition under consideration would not change. (p.82) Since one cannot know that one is either the mug or the lucky player, and since one's credences would change if one were to learn this, these aren't appropriate designators, and so one needn't defer to their conflicting recommendations. Thus we get the right result by means of the following principle:

> *Improved Generalized Reflection Principle*: If a rational agent-at-a-time A respects an agent at-a-time A^* (so designated), and if A^* is an appropriate designator, then A defers over proposition P to A^*.

I think this revised version of the Reflection Principle may be unnecessary to accommodate our problem case. Mahtani acknowledges that the right value, in the case where an agent defers to her future self, should be the mean of the different values she envisions her future self to hold. (p.78) But the acknowledgment that some sort of averaging is required to adjudicate conflicting future opinions should apply just as readily to designators like the mug and lucky. As a first pass, then, one might think that we ought to take the mean of the t_1 credences of the mug, lucky—and whoever else we respect. This would yield the right result without having to appeal to the concept of an appropriate designator, suggesting that the problem Mahtani identifies arises not from the opacity of claims involving co-referring terms, but from abandoning the core commitment of the credence framework to judgements that come in degrees.

One might object that this strategy does not resolve the more fundamental problem, which is that the lack of stability with which lucky and the mug maintain their own credences seems to undermine their standing as experts. (p.83-84) How can their credence functions be expert functions if they would abandon them upon learning who they are? This worry about the standing of lucky and the mug as experts suggests a different approach to our puzzle. If we take seriously that which we have stipulated already—namely, that the mug and lucky know the set-up of the case—this complicates the evidence we can expect them to have at t_1. If I know at t_0 that lucky will get an ace, then the mug knows this as well. Thus at t_1 when the mug turns over her card and learns that she is not lucky, she both has the information that her card reveals, *as well* as the knowledge that lucky's credence in the proposition that *two of the cards dealt out are aces* is different from her own. Moreover, the mug knows (as I do) that lucky is just as worthy of respect at t_1 as she is. Given this, I ought to expect her to internalize the weighting described above: she ought to take into equal account the credences she has after conditionalizing on the information that her card reveals, and what she knows lucky's credences must be after she has done the same. (Similar reasoning applies, of course, to what I ought to expect of lucky. Properly taking into account the fact that I take lucky and the mug to be experts in virtue of knowing everything that I know (and more), then, suggests that we needn't abandon the original generalized reflection principle after all.

3   Practical Implications of the Tenet

Chapter six considers the implications of the tenet for decision theory and welfare economics. Mahtani begins by discussing the *the two-envelope paradox*:

> Suppose that you have before you two envelopes, each of which contains a cheque for some amount of money. You have no idea how much money these two envelopes contain, but you do know that one contains twice as much money as the other. You select an envelope at random. But then you consider: should you stick with the envelope you've selected, or should you switch to the other...Obviously the expected utility of sticking with your current envelope is M (you are guaranteed to get the amount that is in that envelop); whereas if you switch there is a 0.5 chance that you will get double M, and a 0.5 chance that you will get half of M, and so the expected utility of switching is $(0.5)(2M)+(0.5)(0.5M)=1.25M$. Switching, then, seems to have higher expected utility than sticking, and so MEU requires you to switch. But this is a very strange result! Intuitively there is no rational reason for you to switch. And it is not just intuition that speaks against this result, for we can construct a parallel argument in favour of sticking.... (p.111)

The aim of the first part of the chapter is to formulate a restriction on M that rules out as defective the decision table that gives rise to this paradox.

Mahtani begins by entertaining the possibility that M might be a referential term or a definite description. However, she observes that any designator ought to be such that 'we should be certain of the value of whatever the term denotes, such that at every event the term denotes an object of the same value' (p.119-20). While Horgan [2002] takes such 'epistemically rigid designators' to be ones that are stated in terms such that we are certain what those terms denote, Mahtani objects that 'the very idea of knowing what something is seems to fall apart on examination' (p.119). This motivates her to adopt a principle that bypasses this idea:

> *Sameness of expectation of value:* An outcome in a decision table should be designated in such a way that under each event, the agent's expectation of the value of the outcome (so designated) is the same.

I think this refinement of Horgan's strategy gets us the right result in the wrong way. Horgan's solution places a constraint on designators like M that is epistemically virtuous in its own right. It is an epistemically good thing that we know what something is, notwithstanding the difficulty of making sense of this idea. By contrast, in taking our designators to be *whatever* it is that will secure the sameness of expectation of value, one might worry that Mahtani's principle is not motivated independently of the problem that it resolves. This is the aim Mahtani sets for herself at the beginning of the chapter, which she begins by considering the strategy of defining a variable, Z, that stands for the amount of money in the envelope, from which we can reverse engineer the outcomes that yield the right expected utilities. Mahtani rejects this solution on the basis that it fails to explain what gives rise to the paradox; it does little more than stipulate that the designator will be whatever yields the correct values. (p.112) However, in asking us to constrain an

outcome's description by our expectation in its value across events, Mahtani's solution seems vulnerable to the same objection. The distinction between the solution and the values the solution provides collapses in a way that renders the former unable to explain the latter.

The second part of chapter six considers the implications of the tenet for decisions affecting the outcomes for a group of people. Pareto principles, especially *ex ante* Pareto principles, play an important role in debates among utilitarians and egalitarians. However, if credence claims are opaque, an action might be *ex ante* Pareto superior relative to one set of designators, but not another. Mahtani illustrates this with the case where we are forced to choose between giving some medicine to either Alice or Belinda, or splitting it between them. The *ex ante* principle generates conflicting results in the case where 'you are sure that either Ms Smith is Alice and Ms Jones is Belinda, or vice versa, but you don't know which, and your credence is equally divided between these two possibilities' (p.131). In this case, there is no *ex ante* Pareto superior action relative to Alice and Belinda. However splitting is the superior action relative to Ms Smith and Ms Jones. Mahtani's response to this conflict is to adopt the following supervaluationist approach:

> One action is *ex ante* Pareto superior to another *simpliciter*, iff it is *ex ante* Pareto superior relative to *every* set of designators: every set of designators, that is, such that the decision maker is certain that it contains one and only one designator for each of the people concerned. (p.135)

Crucial to Mahtani's argument is the assumption that ignorance be represented as probabilistic indifference over sets of identity claims that partition the space of possibilities. I think there is reason to reject this assumption.

Regarding the assumption of probabilistic indifference, Mahtani writes that 'this is exactly the sort of uncertainty that we should expect given that credence claims are opaque' (p.131). However, this is not the sort of ignorance Tom exemplifies about the proposition that 'George Orwell is Eric Blair'. What makes this, and other similar cases, so compelling is that the identity claim in question seems to have never been considered, much less assigned some probability. While other cases may differ in this regard, the assumption of probabilistic indifference, or of any credence distribution over identity claims, isn't a necessary or distinguishing feature of the general problem case Mahtani seems to have in mind throughout the book. Since a determinate set of values is required to generate the incoherence that motivates the supervaluationist approach, one might worry that this solution isn't responsive to the spirit of the worry Mahtani aims to address.

Even assuming we are on board with modeling ignorance as probabilistic indifference over identity claims, the assumption that these identity claims partition the space of possibilities raises further doubts about Mahtani's solution. It's difficult to accept that we ought to abandon the original *ex ante* principle if we know that there is *some* way that Ms Smith and Ms Jones map onto Alice and Belinda. While Mahtani insists that preferring one set of these designators over the other is arbitrary, one morally significant difference between the two is that only one set of designators can feature in the *description* of the action the agent is deliberating about. As Mahtani notes, the action of giving the medicine to Alice is *not* the action of giving the medicine to Ms Smith, even if these are one and the same individuals. (§6.3) Plausibly, then, what distinguishes Alice and Belinda as the

relevant designators is that they are the individuals whose well-being the agent *intends* to affect in her decision-making. Since any action must be described in terms of one of these pairs of designators, so that there will always be a fact of the matter about who the agent intends to affect—and since the morally relevant facts make it equally permissible to intend to affect the well-being of either Alice and Belinda or Ms Smith and Ms Jones— we can reinterpret the case as involving a brute, but morally permissible intention, rather than an arbitrary choice. This leaves us with a principled way of applying the original *ex ante* principle, whose verdicts coincide with those of the supervaluationist approach.

4   Modifying the Foundations

The last two chapters of the book consider whether the credence framework might be able to accommodate the opacity of credence claims by either (1) rejecting the assumption that states are metaphysically possible worlds, or by (2) rejecting the idea that these states satisfy certain natural requirements. Chapter seven explores the possibility of rejecting these natural requirements, focusing largely on Chalmers' two-dimensionalist approach with its appeal to two sets of possible worlds corresponding to two distinct intensions. The set of worlds that 'verify' an assertion—that ensure that the assertion is *a priori* iff those worlds are actual—is the primary intension of the assertion. The secondary intension of an assertion are the sets of worlds considered counterfactually. According to this framework, worlds where those *a priori* connections we associate with George Orwell don't apply to Eric Blair are worlds where George Orwell is not Eric Blair. And, thus, this approach seems to have a way of explaining the datum that motivates the tenet.

Despite this, Mahtani argues that Chalmers' approach is untenable. The most promising way of operationalizing this approach is to take the primary intension to be determined by the speaker's utterance. But this raises an alignment problem for credal norms since it implies that the objects of credence cannot be shared. The primary intension of 'George Orwell is a writer' might be different depending upon who utters it because different agents may associate different *a priori* connections with that name. This feature makes deference principles nonsensical since it would require one to align one's credence in one object with someone else's credence in a different object. It poses analogous problems for cases of disagreement (where the misalignment is between the object of one's own credence and the different object of the credences of a peer) and conditionalization (where the misalignment is between the different objects of one's own past and present credences). The dilemma, then, is that if we appeal to primary intensions, there is no common object about which we can defer, update, and so on. And, if we appeal to secondary intentions, we cannot accommodate the tenet.

Mahtani considers and rejects several moves one might make in response to this dilemma. While I cannot do justice to the richness of her discussion, I want to suggest that the worry about misalignment may be overstated—or, perhaps, misstated. Mahtani maintains that conditionalization will require that my credences at t_1 be constrained by my credences at t_0 even where, for instance, it is *a priori* for me at t_0 that George Orwell is the person whose photograph hangs in the department's library above that name, though this fails to be *a priori* for me at t_1 after I have learned more about George Orwell's life history. However, it's unclear that a version of conditionalization formulated in terms of centred worlds, which is what Mahtani envisions, would yield this result rather than

simply fall silent in cases where, by stipulation, the objects in question were different—a perfectly acceptable result given that one would not expect such cases to arise very often. Since Mahtani does not formulate the updating rule she has in mind, however, one can only speculate about this.

Finally, chapter eight explores the possibility of rejecting the assumption that states are metaphysically possible worlds. Mahtani considers the possibility that states might be more fine-grained worlds, which she thinks are best understood as sets of sentences in some natural or artificial language. She concludes by briefly considering the possibility that the objects of credence might be structured propositions instead. Mahtani lays out the challenges for each of these views. She also considers the difficulties that arise when we attempt to impose constraints of completeness and coherence on the sets of worlds the first set of these views involve. I will end with a brief observation about this last part of the discussion. Mahtani notes that the assumption that the nature of states is up for grabs poses problems for dutch book arguments and other justifications for probabilism since one might resist the conclusion of such arguments simply by adjusting one's interpretation of states. As she puts it:

> For the argument to work, we ought to have an independent account of the nature of states...But in fact our judgements about rationality help mould our interpretation of states in the first place, and the conclusion of a dutch-book argument will simply reflect these judgements. (p.192)

There are different lessons one might draw from this observation. I want to suggest again that the right thing to say about this is that interpreting the credence framework as capable of these sorts of adjustments misrepresents the way it is meant to be understood. The norms that govern the framework within which identity claims are assumed ought to be regarded as distinct from the norms on linguistic competence that govern these identity claims. It should come as no surprise, then, that one set of these norms must be in place in order to evaluate the other.

Despite these issues, Mahtani offers us a rich and important book. It is required reading for those working at the intersection of the philosophy of language and formal epistemology, and all users of the credence framework can benefit from its lessons.

LISA CASSELL
University of Maryland, Baltimore County, USA
lisarcassell@gmail.com

References

Terry Horgan. "the two-envelope paradox, nonstandard expected utility, and the intensionality of probability.". *Nous*, 34:578–603, 2002.