Penultimate draft. Final version to appear in *The Review of Symbolic Logic*, 2018

BELIEVING EPISTEMIC CONTRADICTIONS

BOB BEDDOR AND SIMON GOLDSTEIN

Abstract. What is it to believe something might be the case? We develop a puzzle that creates difficulties for standard answers to this question. We go on to propose our own solution, which integrates a Bayesian approach to belief with a dynamic semantics for epistemic modals. After showing how our account solves the puzzle, we explore a surprising consequence: virtually all of our beliefs about what might be the case provide counterexamples to the view that rational belief is closed under logical implication.

§1. The Puzzle. Ari the burglar has been casing the house for hours. As far as she can tell, not a mouse is stirring. Consequently, Ari believes the house is empty.

Still, Ari is an experienced burglar; she knows that even the most thorough reconnaissance is fallible. Thus she allows that there's some possibility—albeit very remote—that an inconspicuous resident is still inside.

Given this, it seems we should be able to report Ari's belief state as follows:

(1) # Ari believes the house is empty and might not be empty.

Key words and phrases. Epistemic modals, Dynamic semantics, Probability, Belief Reports.

Special thanks to Johan van Bentham, Justin Bledin, Fabrizio Cariani, Cat Saint Croix, Andy Egan, Thony Gillies, Jeff King, Ernie Lepore, Hanti Lin, Salvador Mascarenhas, Mike Raven, Jonathan Schaffer, Alex Worsnip, Seth Yalcin, and two anonymous referees at *The Review of Symbolic Logic* for detailed comments. Thanks also to audiences at Oxford University, Rutgers University, NYU, KU Leuven, the 2016 Central APA, the 2016 Semantics and Pragmatics Workshop at HSE, and the 17^{th} Formal Epistemology Workshop for helpful discussion.

Meeting
Edited by Unknown

But (1) sounds odd. In uttering (1), it sounds like you're attributing to Ari an incoherent belief.[1] This is surprising, since nothing in our set-up seemed to saddle Ari with incoherence. On the face of it, Ari is being perfectly rational: forming beliefs in response to her evidence, without being overly confident in her conclusions.

We have here the makings of a puzzle, which can be formulated in more general terms. It's a sad but indisputable fact that humans are fallible: we've all held beliefs that turned out to be mistaken. Recognizing this fact, it seems rational for an agent to hold some belief, while also acknowledging that it's possible that this belief is mistaken. At the very least, it seems that this stance is *coherent*. Letting '$\Diamond$' represent epistemic possibility, we can formulate this principle as follows:

FALLIBILITY: It's sometimes coherent for an agent to believe ϕ and also believe $\Diamond\neg\phi$.

For example, Ari believes that the house is empty, while also believing that there's some possibility that someone is still inside.

While Fallibility seems plausible, our reaction to (1) motivates a principle that stands in tension with it. Following Yalcin [2007], let us refer to any sentence of the form $\phi \wedge \Diamond\neg\phi$ as an *epistemic contradiction*. Given the oddity of (1), it's tempting to maintain that it is always incoherent to believe an epistemic contradiction:

NO CONTRADICTIONS: It's never coherent to believe $(\phi \wedge \Diamond\neg\phi)$.

But it is *prima facie* difficult to reconcile these two principles. Is any story about beliefs involving epistemic modals up to the task?

§2. Roadmap. Whenever we encounter two principles that stand in tension with each other, it is prudent to examine whether the principles are really all that plausible. We thus start by defending our principles: in §3, we argue that both are well-supported by linguistic data. Consequently we should solve our puzzle not by denying one of our principles, but rather by providing a semantics for epistemic modals and belief reports that validates both of them, thereby dissolving the appearance of inconsistency.

This is no trivial task. The classical semantics for epistemic modals doesn't deliver the desired results. As we show in §4, standard contextualist (Kratzer [1981], [2012]) and relativist (Egan [2007]; Stephenson [2007a],

[1]It's important that both conjuncts of the complement clause (*the house is empty and (it) might not be empty*) scope under *believes*. The sentence:

(i) Ari believes [the house is empty]; and it might not be empty.

does not ascribe to Ari an incoherent belief. We discuss versions of (i) in §3.

[2007b]; MacFarlane [2011]) approaches to modals are forced to give up either Fallibility or No Contradictions.

By itself, this may strike some as unsurprising. After all, Yalcin [2007], [2011] has argued that the classical semantics has trouble explaining the infelicity of epistemic contradictions in embedded contexts. However, we show that while leading non-classical accounts of beliefs involving epistemic modals (Veltman [1996]; Gillies [2001]; Yalcin [2007], [2011], [2012b]; Willer [2013]) validate No Contradictions, they invalidate Fallibility (§5). As a result, they also fail to provide a satisfactory solution to our puzzle.

In §6 we present our own solution, which integrates a Bayesian approach to belief with a dynamic semantics for epistemic modals. According to our theory, A believes ϕ iff A assigns a sufficiently high credence to the result of updating A's information with ϕ. Assuming a dynamic account of updating, this theory validates both of our principles, thereby solving the puzzle. In broad outline, the solution is this. Our theory borrows from Bayesianism the insight that belief requires sufficiently high credence, not certainty. At the same time, the dynamic account of updating predicts that modal beliefs are 'transparent': to believe that the house might not be empty is to be less than certain that it's empty. And so our theory predicts that whenever an agent believes ϕ without being certain of it, she counts as believing $\Diamond\neg\phi$, thereby validating Fallibility. However, the dynamic account of updating also predicts that no body of information can coherently be updated with $\phi \wedge \Diamond\neg\phi$. Hence no coherent agent will believe an epistemic contradiction, thereby validating No Contradictions.

After developing our solution to the puzzle, we defend and extend it. §7 explores a surprising consequence of our account of belief: coherent agents can believe ϕ and believe $\Diamond\neg\phi$, but they cannot coherently believe an obvious entailment of these beliefs: $\phi \wedge \Diamond\neg\phi$. Thus our solution entails that rational belief is not closed under logical implication. Indeed, our view suggests that counterexamples to closure occur all the time—no need to look to lotteries and prefaces. Finally, §8 defends our solution from objections.

§3. Defending our Principles.

3.1. Fallibility. Let us start with Fallibility—the thesis that it's sometimes coherent to believe ϕ and to also believe $\Diamond\neg\phi$. In addition to its intuitive plausibility, we offer three arguments in support of this thesis.

The first is the argument from *concessive belief attributions.* Let a concessive belief attribution (CBA) be a discourse of the form:

(2) I believe ϕ. But $\Diamond\neg\phi$.

Such discourses seem perfectly felicitous. Consider, for instance:

(3) I believe/think the movie starts at 7, but $\left\{ \begin{array}{l} \text{it might start later} \\ \text{I might be mistaken} \end{array} \right\}$.[2]

Presumably it is infelicitous to make an assertion if it is impossible for a coherent agent to have a true belief in its content. But if Fallibility were false, no coherent agent could truly believe a CBA. After all, suppose the first conjunct of a CBA (*I believe* ϕ) is true. If Fallibility were false, the second conjunct (*But* $\Diamond\neg\phi$) could not be coherently believed. More generally, if Fallibility were false, then anyone who uttered a CBA would be committed to having an incoherent doxastic state.

The second argument for Fallibility comes from considering discourses of the form:

(4) Ari believes/thinks the house is empty. But she realizes/recognizes that it might not be.

(4) sounds perfectly fine.[3] But presumably both realizing and recognizing entail believing. So if Fallibility were false, (4) would ascribe to Ari an incoherent doxastic state.

Our third argument for Fallibility is more theoretical. According to this argument, Fallibility follows from two further principles, both of which are supported by independent data. The first principle is that it's coherent to hold uncertain beliefs:

Uncertain Belief: It's possible to coherently believe ϕ without being certain that ϕ.

The idea that belief doesn't require certainty is widely taken for granted.[4] It's also well-supported by our everyday *belief* and *certainty* talk. Some examples:

(5) I believe, but am not certain, that the land was granted to Mr. Baca for pasturing purposes...[5]

(6) I believe but am not certain that Palestine is not a party to the statute of the International Court of Justice, and I believe but am not certain that the General Assembly has not addressed an invitation to Palestine to join the Convention.[6]

---

[2] Concessive knowledge attributions (that is, sentences of the form, *I know* ϕ*, but* $\Diamond\neg\phi$) have attracted a great deal of attention in recent years. Interestingly, while it's frequently observed that concessive knowledge attributions are infelicitous (see, e.g., Lewis [1996]; Rysiew [2001]; Stanley [2005]; Dodd [2011]; Worsnip [2015]), the comparative felicity of CBAs seems to have gone unnoticed.

[3] In a similar vein, Hawthorne, Rothschild, and Spectre [2016] observe that it sounds coherent to say, *I believe it's raining, but I know it might not be* (1396).

[4] Though see Clarke [2013] and Dodd [forthcoming] for dissent.

[5] *Executive Documents of the House of Representatives*, vol.125-126: 40.

[6] `https://lettersblogatory.com/2015/01/02/palestine-signs-new-york-convention/`

Such utterances seem to express coherent doxastic states.[7]

In case the reader is tempted to attribute the coherence of (5) and (6) to some special feature of first-person belief reports, note that it is also natural to ascribe uncertain beliefs to others. For example:

(7) Ari believes/thinks that the house is empty, but she's not certain of it.

(7) seems to ascribe a perfectly coherent doxastic state to Ari.

The second step in the argument is to posit a connection between uncertainty and believing possible:

Uncertainty-Possibility Link: If an agent A is coherent, then if A isn't certain that ϕ, A is in a position to believe $\Diamond\neg\phi$.

We suspect that Uncertainty-Possibility Link will prove more controversial than Uncertain Belief. A couple of clarificatory remarks may help forestall some immediate objections. First, in order for the principle to be remotely plausible, we need to interpret it as making a claim about *epistemic* possibility, rather than physical or metaphysical possibility. Second, the 'in a position' qualification here is important. To see this, consider an agent who is not certain that ϕ, but has never even considered the question of whether ϕ. Arguably, it is at least somewhat counterintuitive to say that they believe that ϕ might be false. Thus, Uncertainty-Possibility Link should be interpreted as making the following claim: if a coherent agent isn't certain that ϕ, they are committed to believing (perhaps on further reflection) that it's epistemically possible that ϕ might be false.

Even with these clarifications, some may deem Uncertainty-Possibility Link implausible. After all, it's widely held that certainty is a very demanding state. For example, many philosophers maintain that if one is certain of ϕ, one should be willing to accept a bet where one wins a penny if ϕ is true, and one loses life and limb otherwise. Presumably, very few ordinary beliefs are held with this degree of conviction.[8] But if we are almost never in a position to believe an ordinary proposition ϕ with certainty, Uncertainty-Possibility Link entails that we are almost always in a position to believe $\Diamond\neg\phi$, and hence almost never in a position to believe $\neg\Diamond\neg\phi$. But this seems wrong. In ordinary contexts, it would be natural to say:

(8) Jim believes there's no possibility the Lions will win.

even if one doesn't think that Jim would stake his life on the Lions losing.

[7]See Christensen [2005]: 21 and Hawthorne, Rothschild, and Spectre [2016]: 1395 for similar observations.

[8]For the view that certainty is rarely (if ever) attained, see Russell [1912]; Unger [1975], among others.

However, proponents of Uncertainty-Possibility Link have a natural response. Observe that in ordinary contexts, it would be equally natural to say:

(9) Jim is certain that the Lions will lose.

This suggests that the truth-values of certainty ascriptions differ from context to context, where here the context could be either that of the speaker or that of the subject of ascription.[9] According to this view, only in the most demanding contexts will a claim of the form *A is certain that* ϕ entail that A is willing to bet their life on ϕ. In most ordinary contexts, less is required.[10]

Indeed, reflecting on the assertability conditions of certainty ascriptions provides a compelling argument for Uncertainty-Possibility Link. Consider:

(10) a. Ari isn't certain that the house is empty.
b. ?? But she doesn't believe/think there's any possibility there's someone inside.

Assuming we're in a context in which Ari has considered the question of whether the house is empty, following (10a) with (10b) sounds very odd. Uncertainty-Possibility Link explains this: (10) ascribes an incoherent doxastic state to Ari.[11]

Thus, both Uncertain Belief and Uncertainty-Possibility Link are well-supported by the data. And their conjunction entails Fallibility. After all, by Uncertain Belief it's coherent for an agent to believe ϕ without being certain of ϕ. By Uncertainty-Possibility Link, such an agent will always be in a position to believe $\Diamond\neg\phi$. *A fortiori*, it will sometimes be coherent for an agent to believe both ϕ and $\Diamond\neg\phi$, as Fallibility maintains.

We have presented three arguments for Fallibility: the argument from CBAs, the argument from the *realize-belief* entailment, and the argument

[9] See Lewis [1979]; Stanley [2008]; Beddor [2016]. Note that if certainty ascriptions are context-sensitive, then it would be more accurate to formulate Uncertainty-Possibility Link in the formal mode: if *A isn't certain that* ϕ is true in a context, then *A is in a position to believe* $\Diamond\neg\phi$ will also be true in that context.

[10] Some authors will resist the conclusion that certainty attributions are context-sensitive. These authors will insist that the speaker's utterance of (9) is "loose talk": false but pragmatically acceptable (Unger [1975]; Lasersohn [1999]; Kennedy [2007]). Positing such rampant falsity in ordinary conversation strikes us as a cost. That said, the "loose talk" approach is still compatible with Uncertainty-Possibility Link, provided it is reformulated in terms of assertability conditions: if *A isn't certain that* ϕ is assertable in a context, then *A is in a position to believe* $\Diamond\neg\phi$ will also be assertable in that context.

[11] One potential complication for this argument is that *believes* is neg-raising, and so perhaps the logical form of *A doesn't believe* ϕ is really, *A believes* $\neg\phi$. To control for this, just rewrite the sentence with the quantifier, *nobody*. To our ears, *Nobody is certain that the house is empty, but nobody thinks there's any possibility there's someone inside* is equally marked.

from the conjunction of Uncertain Belief and Uncertainty-Possibility Link. Of course, not all readers will be fully persuaded by these arguments; we will discuss some objections to Fallibility in §8. But for now, we propose to accept Fallibility as a working hypothesis.

3.2. No Contradictions. Turn now to No Contradictions. We've already provided the main argument for this principle: (1) (*Ari believes the house is empty and might not be empty*) seems to ascribe an incoherent doxastic state to Ari. What's more, nothing hinges on the details of the example. As far as we can tell, every instance of the schema:

(11) # A believes $(\phi \wedge \Diamond\neg\phi)$.

seems to ascribe incoherent beliefs to A.

We can strengthen the case for No Contradictions by observing that the oddity of (11) is not an isolated phenomenon. Epistemic contradictions sound incoherent in a variety of contexts. It's often been noted that assertions of unembedded epistemic contradictions sound bizarre:

(12) # The house is empty and might not be empty.[12]

Various authors have also noted that epistemic contradictions sound odd in the antecedents of indicative conditionals, as well as under other attitude verbs—for instance, *suppose* and *imagine* (Yalcin [2007], [2011]; Anand and Hacquard [2013]; Dorr and Hawthorne [2013]). This suggests that there is a genuine and general phenomenon here.[13]

Some readers may be inclined to concede instances of (11) are infelicitous, but doubt that this is best explained by the hypothesis that (11) ascribes incoherent beliefs to agents. But then what explains the infelicity of (11), if not the incoherence of the underlying doxastic state? Providing an explanation is no easy matter. Indeed, the puzzle can be reframed in a way that relies directly on our linguistic judgments, rather than No Contradictions. What account of beliefs involving epistemic modals will validate Fallibility, while also explaining the infelicity of (11)?

Having motivated our two principles, we now turn to consider whether any account of beliefs involving epistemic modals can validate both of them.

§4. Troubles for the Classical Semantics. Orthodoxy has it that modals quantify over possibilities (Kratzer [1981], [2012]). But they don't quantify over just any possibilities. Their domain is restricted by a contextually determined set of worlds: *the modal base*. Possibility modals such

---

[12]See Veltman [1996]; Gillies [2001]; Yalcin [2007], [2011]; Willer [2013].

[13]Note that both embedded and unembedded epistemic contradictions remain infelicitous when a contrast marker such as *but* is used. For instance, *Ari believes the house is empty but might not be* is quite bizarre (even if it is slightly less bizarre than (1)).

as *might* existentially quantify over the modal base; necessity modals such as *must* universally quantify over it. Let c be a context, w an index, $[\![\phi]\!]^c$ the set of indices w such that ϕ is true at $\langle c, w \rangle$, and B_c^w the modal base determined by c and w. Contextualists propose:

Contextualist Might: $[\![\Diamond\phi]\!]^c = \{w \mid B_c^w \cap [\![\phi]\!]^c \neq \emptyset\}$.

Epistemic modals are evaluated using an *epistemic modal base*—a set of worlds reflecting the epistemic state of some contextually determined agents. To illustrate, a sentence such as:

(13) The house might not be empty.

is true iff the prejacent (*The house is not empty*) is compatible with the epistemic state of the contextually determined group.[14]

Standard relativist accounts of epistemic modals (Egan [2007], Stephenson [2007a], [2007b]; MacFarlane [2011]) are similar. Their main point of departure is that they take the epistemic modal base to reflect the epistemic state of an assessor—an individual who is interpreting the modal. Let a be a context of assessment and B_c^a the c, a-determined modal base. Relativists propose:

Relativist Might: $[\![\Diamond\phi]\!]^c = \{\langle w, a \rangle \mid B_c^a \cap [\![\phi]\!]^c \neq \emptyset\}$.

This allows that an utterance of (13) could be true relative to one context of assessment (where the relevant folks' epistemic state leaves open the possibility that the house is occupied), and false relative to a different context of assessment (where the relevant folks' epistemic state entails the house is empty).

In order for the classical semantics to generate predictions about our puzzle, we need to say more about the nature of the epistemic modal base. In what follows, we look at the two most natural options: the *knowledge-based approach* and the *belief-based approach*. We argue that neither validates both of our principles; hence neither provides a satisfactory resolution to our puzzle.

According to the knowledge-based approach, an epistemic modal base is the set of possibilities compatible with what the relevant agents know, or can come to know (Hacking [1967]; Kratzer [1981], [2012]; DeRose [1991]; Egan, Hawthorne, and Weatherson [2005]; Stanley [2005]; Stephenson [2007a]; Hawthorne [2007], [2012]; Egan and Weatherson [2011]). On this approach, if A believes $(\phi \wedge \Diamond\neg\phi)$, then what A believes is equivalent to:

[14] Our exposition makes a standard simplification by omitting the role that the ordering source plays in most versions of a Kratzerian semantics. We also omit the fact that some take epistemic modals to convey indirectness (see e.g., Fintel and Gillies [2010]).

(14) ϕ and ($\neg\phi$ is compatible with what the relevant agents know).

But why would this be incoherent? Suppose A is the only relevant agent. Since belief doesn't entail knowledge, A could believe ϕ without knowing ϕ. What's more, A could believe truly she's in such a position—that is, she could believe truly that she believes ϕ without knowing ϕ.[15] Thus the knowledge-based approach fails to predict No Contradictions.

Some might think that this is too quick. Perhaps even though a subject can believe ϕ while failing to know ϕ, no subject can coherently take herself to believe ϕ while failing to know ϕ. One way of motivating this would be to appeal to the idea that knowledge is the norm of belief (Williamson [2000]: 47; Sutton [2005], [2007]; Bird [2007]; Huemer [2007]).[16]

But we find this strategy unconvincing. There are certainly agents who take themselves to hold beliefs that don't amount to knowledge. Consider Thelma the theist, who professes to believe that God exists, while also claiming that she doesn't know that God exists: it's a matter of faith. Or consider Louise the lottery ticket holder, who believes her lottery ticket will lose (on statistical grounds), but also claims not to know it will lose, on the grounds that knowledge requires safety (McGlynn [2013]). It's natural to describe their doxastic states thus:

(15) ✓ Thelma believes that God exists and that she doesn't know God exists.

(16) ✓ Louise believes that her ticket will lose and that she doesn't know her ticket will lose.

But the knowledge-based approach predicts that (15) and (16) are equivalent to:

(17) # Thelma believes God exists and might not exist.

(18) # Louise believes her ticket will lose and might win.

So the knowledge-based approach does not explain the difference in felicity between these pairs of sentences.[17]

---

[15] If there are relevant agents besides A, we expect cases in which one can truly believe an epistemic contradiction to be even more common: they'll include any case in which A believes ϕ and also believes that the other relevant folks don't know ϕ.

[16] In a similar vein, Dorr and Hawthorne [2013] suggest that some uses of *believes* implicate that the believer believes that she knows the complement clause (910, n.60).

[17] While our main criticism of the knowledge-based approach is that it doesn't validate No Contradictions, it's also unclear whether it validates Uncertainty-Possibility Link. The only way that the knowledge-based approach could validate this principle is if knowing ϕ entails being in a position to be certain that ϕ. While some authors endorse this entailment (Ayer [1936]; Moore [1959]; Unger [1975]), cases such as Radford's unconfident examinee provide grounds for doubt (Radford [1966]). Intuitively, the examinee knows the answer to the examiner's question, even though he isn't in a position to be certain of it (Armstrong [1973]; Stanley [2008]; McGlynn [2014]; Beddor [2016]).

Given these difficulties for the knowledge-based approach, one might adopt a belief-based approach, according to which an epistemic modal base is the set of possibilities compatible with what the relevant agents believe. This approach is in a better position to capture No Contradictions. According to the belief-based approach, if A believes $(\phi \wedge \Diamond\neg\phi)$, then what A believes is equivalent to:

(19) ϕ and ($\neg\phi$ is compatible with what the relevant agents believe).

If we assume that the relevant folks include A, this entails that A is committed to believing a Moore-paradoxical proposition:

(20) ϕ and I don't believe ϕ.

It's a familiar observation that such Moorean beliefs seem incoherent. And so the incoherence of believing an epistemic contradiction is explained in terms of the incoherence of Moorean belief.[18]

However, giving this explanation of No Contradictions requires giving up Fallibility. Here's why. Suppose Fallibility holds, and hence that there's a coherent agent who believes ϕ and also believes $\Diamond\neg\phi$. For concreteness, let's focus on Ari, who believes the house is empty, but also believes there's a possibility that someone is inside. Given the belief-based approach, it follows that Ari believes that it's compatible with what she believes that the house isn't empty. And so Ari is committed to having a Moorean belief: *The house is empty and I don't believe the house is empty.* But then Ari is just as incoherent as someone who believes an epistemic contradiction. Thus the belief-based approach only vindicates No Contradictions at the expense of giving up Fallibility.

At this point some may question the way we have argued against the classical semantics. We've been assuming that the epistemic modal base is *either* the set of possibilities compatible with what the relevant agents know or the set of possibilities compatible with what the relevant agents believe. But why assume that there is a context-invariant answer to this question? In some contexts it may be the former, in others, the latter. And perhaps sometimes it's simply indeterminate.[19]

While this seems reasonable, it doesn't help the classical semantics evade our challenge. After all, one candidate for the epistemic modal base is the set of possibilities compatible with what the relevant agents know. This predicts that there should be coherent readings of belief reports embedding epistemic contradictions. That is, there should be an available reading of

[18] By now there is a large literature on why such Moorean beliefs are absurd. For discussion, see Hintikka [1962]; Williams [1994]; Almeida [2001]; Green and Williams [2007]; Holliday and Icard [2010], among many others.

[19] This proposal would fit naturally with 'flexible contextualism' (Dowell [2011]), according to which the community of relevant agents varies with the context of utterance.

(15) on which it's equivalent to (17), and an available reading of (16) on which it's equivalent to (18). Indeed, we'd expect listeners to converge on these readings, due to general principles of charity. However, as we have seen, there's no readily available coherent reading of belief reports embedding epistemic contradictions. Those who leave it to context to determine the epistemic state in question will have trouble explaining this observation.

§5. Troubles for Non-Classical Semantics. For some readers, it may come as no surprise that the classical semantics has trouble explaining the infelicity of epistemic contradictions in belief reports. Yalcin [2007], [2011] has shown that epistemic contradictions are infelicitous when embedded under various operators, which poses a challenge to the classical semantics. Compare:

(21) # Suppose it's raining and it might not be raining.

(22) ✓ Suppose it's raining and I don't know/believe it's raining.

One might hope for a unified treatment of epistemic contradictions in embedded contexts: whatever explains the infelicity of (21) also explains the infelicity of (1).

Indeed, a variety of non-classical semantics for *might* explain the infelicity of epistemic contradictions in embedded contexts. In doing so, they validate No Contradictions. However, we show that these theories are forced to deny Fallibility. For reasons of space, we focus on one implementation of a non-classical semantics for *might*: a version of the update semantics discussed in Veltman [1996]; Gillies [2001]; Yalcin [2012a], [2012b]; and Willer [2013]. However, the problem we raise generalizes to other non-classical semantics, such as the static semantics developed by Yalcin [2007], [2011] and Moss [2015].

Update semantics is a type of dynamic semantics. In a dynamic semantics, the meaning of an expression is not its truth conditions. Rather, the meaning of an expression is its *context change potential.* This *ccp* is a function that takes as input a context and returns the result of updating that context with the expression. In a slogan: the meaning of an expression is its ability to change a body of information.[20]

Consider a language $\mathscr{L}$ containing a set of atomic sentences $\{\alpha_1, ..., \alpha_n\}$ closed under *might* ($\Diamond$), *and* ($\wedge$), and *not* ($\neg$). Let a possible world w be a function from atomic sentences to truth values. Let a context s be a set of possible worlds. According to update semantics, the interpretation of $\mathscr{L}$ is a function $[\cdot]$ from sentences in $\mathscr{L}$ to ccps, functions from contexts to contexts, defined recursively as follows:

[20]For important contributions to the dynamic tradition, see Stalnaker [1973]; Karttunen [1974]; Heim [1982], [1983]; Groenendijk and Stokhof [1990], [1991a].

Update Semantics:

1. $s[\alpha] = \{w \in s \mid w(\alpha) = 1\}$
2. $s[\phi \wedge \psi] = s[\phi][\psi]$
3. $s[\neg\phi] = s - s[\phi]$
4. $s[\Diamond\phi] = \{w \in s \mid s[\phi] \neq \emptyset\}$

According to this semantics, an atomic sentence narrows down a context to the worlds where it is true. A conjunction affects the context in two steps: first, it updates the context with the first conjunct; next, the resulting context is updated with the second conjunct. The negation of a sentence ϕ updates the context with the $\neg\phi$ worlds.

These first three clauses are updates: they affect the context by narrowing down the possible worlds in it. By contrast, *might* is a test. Rather than narrowing down the worlds in a context, $\Diamond\phi$ checks whether the context is compatible with ϕ. For example, (13) (*The house might not be empty*) will leave the context unchanged provided there's at least one world in the context where the house isn't empty. Otherwise, the context crashes (represented as the empty set of worlds $\emptyset$).

This semantics predicts that unembedded epistemic contradictions are infelicitous (Veltman [1996]; Gillies [2001]). To see this, let us first define a notion of inconsistency for Update Semantics. A sentence ϕ is inconsistent just in case updating with ϕ is guaranteed to crash any context:

Consistency: A sentence ϕ is consistent iff $\exists s$: $s[\phi] \neq \emptyset$.

It is easy to show that epistemic contradictions are inconsistent in Update Semantics:

Fact 1 (Epistemic Contradictions are Inconsistent). For any descriptive (non-modal) sentence ϕ and any context s: $s[\phi \wedge \Diamond\neg\phi] = \emptyset$.

Proof. $s[\phi \wedge \Diamond\neg\phi] = s[\phi][\Diamond\neg\phi]$. When ϕ is descriptive, $s[\phi]$ only contains ϕ worlds. So this set always fails the test performed by $\Diamond\neg\phi$. So $s[\phi \wedge \Diamond\neg\phi] = \emptyset$. ⊣

To illustrate with an example, take (12) (*The house is empty and it might not be empty*). Suppose our context s contains three worlds: w and u, in which the house is empty, and v, in which there's someone inside. The first conjunct (*The house is empty*) narrows the context down to worlds where the house is empty, giving us $\{w, u\}$. The second conjunct then tests to see whether this updated context contains any worlds where the house isn't empty. Since the updated context fails this test, the sentence crashes. (See Figure 1.)

However, thus far Update Semantics doesn't make any predictions about our puzzle, because it lacks a semantics for *believes*. Let's now enrich $\mathscr{L}$

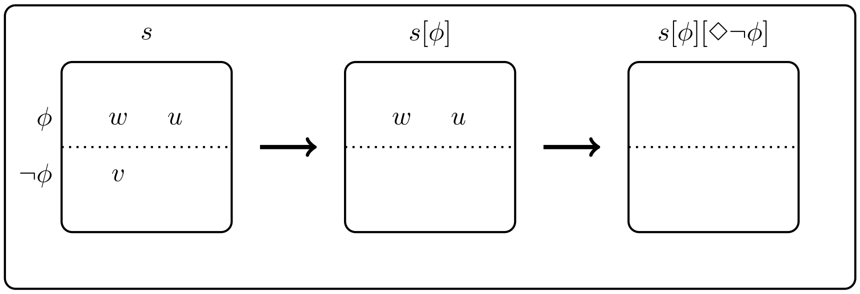

FIGURE 1. Updating with $\phi \wedge \Diamond\neg\phi$

with a *believes* operator (B_A). The standard dynamic semantics for *believes* analyzes belief in terms of *support*, where support is defined as a fixed point:

SUPPORT: A context s supports ϕ ($s \models \phi$) iff $s[\phi] = s$.

On this definition, a context supports a sentence iff updating the context with that sentence has no effect on the context. According to the standard account, an agent then believes ϕ iff her doxastic state supports ϕ.[21] More precisely, suppose that an agent A's doxastic state at a world w is characterized by a set of worlds (s_A^w): these are the worlds consistent with what A believes at w. Then:

BELIEF AS SUPPORT: $s[B_A\phi] = \{w \in s \mid s_A^w \models \phi\}$.

To illustrate, consider:

(23) Ari believes the house is empty.

According to Belief as Support, (23) narrows down a context to those worlds where Ari's doxastic state supports *The house is empty*, which in turn obtains iff there is no world in her doxastic state in which the house is occupied.

This approach has the advantage of validating No Contradictions. By Fact 1, epistemic contradictions are semantically inconsistent. And so no non-empty set of worlds supports an epistemic contradiction. *A fortiori*, no coherent agent's doxastic state supports an epistemic contradiction. By Belief as Support, it follows that no coherent agent believes an epistemic contradiction.[22]

[21] This account was first proposed by Hans Kamp, and is defended in Heim [1992]; Zeevat [1992]; Yalcin [2012a], [2012b]; Willer [2013].

[22] This strategy generalizes smoothly to explain the infelicity of epistemic contradictions under other attitude verbs. Take *supposes*. Update Semanticists can hold that *A supposes* ϕ narrows down the context to those worlds where A's suppositional state

Unfortunately, Belief as Support invalidates Fallibility. To see this, consider:

(24) Ari believes the house might not be empty.

Given Update Semantics and Belief as Support, (24) updates the context with the information that Ari's doxastic state contains at least one world where the house is occupied. But this means that (23) is false. More generally, $B_A\phi$ and $B_A\Diamond\neg\phi$ provide incompatible instructions for updating the context. And so combining Update Semantics with Belief as Support forces us to abandon Fallibility.[23]

Let's take stock. We've argued that the leading accounts of what it is to believe that something might be the case fail to satisfactorily resolve our puzzle. Specifically, we've canvassed two versions of a classical semantics for epistemic modals (the knowledge-based approach and the belief-based approach) as well as a standard dynamic approach (Update Semantics conjoined with Belief as Support). The knowledge-based approach failed to validate No Contradictions, whereas both the belief-based approach and the dynamic approach invalidated Fallibility.

Is there an alternative approach that can validate both Fallibility and No Contradictions? We think so. In what follows we present our own solution, which integrates Update Semantics with a Bayesian account of belief. We show that this position validates all of our principles, thereby resolving the puzzle.

§6. A New Semantics for Belief Reports. Here, in a nutshell, is our proposal. *An agent believes ϕ iff she assigns a sufficiently high credence to the result of updating her information with ϕ.* This proposal synthesizes Update Semantics with a 'Lockean' account of belief, according to which

supports ϕ. More precisely, let sup_A^w be the worlds compatible with what A supposes at w, and let Su_A abbreviate *A supposes*. Then: $s[Su_A\phi] = \{w \in s \mid sup_A^w \models \phi\}$. Since no set of worlds supports an epistemic contradiction, epistemic contradictions cannot be supposed: (21) is always false. (Cf. the static treatment of *supposes* in Yalcin [2007].)

[23]Other non-classical semantics for epistemic modals, such as those developed by Yalcin [2007], [2011] and Moss [2015], arrive at much the same impasse. For example, Yalcin relativizes the truth conditions of sentences to both a world w and an information state s. (As in Update Semantics, s is a set of worlds.) On Yalcin's semantics, $\Diamond\phi$ is true at some w, s iff s contains at least one world where ϕ is true. Yalcin combines this with a semantics for belief reports according to which *believes* shifts the information state to the believer's doxastic alternatives: $B_A\phi$ is true at some w, s iff for every world w' in s_A^w, ϕ is true at w', s_A^w. This predicts that (23) and (24) have incompatible truth conditions: (23) is true iff the house is empty at every world in Ari's doxastic state, and (24) is true iff the house is occupied at some world in Ari's doxastic state. So, much like Belief as Support, Yalcin's treatment of modal beliefs invalidates Fallibility.

belief reduces to sufficiently high credence.[24] The resulting synthesis combines the primary advantages of both approaches, thereby resolving our puzzle. From Update Semantics, we borrow the resources to validate both No Contradictions and Uncertainty-Possibility Link. From the Lockean view, we borrow the resources to validate Uncertain Belief. And by validating both Uncertainty-Possibility Link and Uncertain Belief, we thereby validate Fallibility.

To introduce the details of our proposal, it may help to start with a simple version of the Lockean view, according to which A believes ϕ iff A assigns a sufficiently high credence to the set of ϕ worlds. Let $[\![\cdot]\!]$ assign to each descriptive sentence of the language the set of worlds where it is true. Let Pr^w_A be A's credence function at w, and let t denote some threshold between 0 and 1. Lockeans propose:

Lockean Belief: $[\![B_A\phi]\!] = \{w \mid Pr^w_A([\![\phi]\!]) > t\}$

That is, $B_A\phi$ is true iff A assigns a credence greater than t to the set of worlds where ϕ is true.

While we'll argue shortly that Lockean Belief requires revision, the basic idea behind the Lockean approach holds considerable appeal. Unlike Belief as Support, Lockean Belief sheds light on the connection between outright belief and degrees of belief. In particular, it validates plausible inference patterns linking these two notions, for instance:

(25) a. Fred believes it's raining. $\Rightarrow$
b. Fred is fairly/quite confident that it's raining.

In addition, Lockean Belief validates Uncertain Belief. After all, it's coherent to have a high credence in ϕ without being certain that ϕ. According to Lockean Belief, a high credence is all that's required for believing outright.

Despite its attractions, Lockean Belief does not solve our puzzle. Taken by itself, it does not validate either Fallibility or No Contradictions. Indeed, taken by itself, it is not clear how it accounts for beliefs about what might be the case. This lacuna is particularly evident if we adopt the dynamic approach to epistemic modals from §5. After all, on the dynamic approach there is no set of worlds where $\Diamond\phi$ is true. Thus the task that now faces us is to extend Lockean Belief to modal beliefs in a way that validates our two principles.

[24]For defenses of a Lockean view, see Foley [1993], [2009]; Christensen [2005]; Sturgeon [2008]; Dorst [forthcoming]. Note that we will be exploring a semantic version of the Lockean position, according to which a belief ascription of the form, $B_A\phi$ simply means than that A has a sufficiently high credence in ϕ. This semantic version of Lockeanism has not been widely discussed, though see Pettigrew [2015], which argues that those inclined towards both accuracy dominance arguments for probabilism and Lockeanism would do well to embrace an analytic version of the latter.

To do this, we propose a dynamic twist to Lockean Belief. On a dynamic approach, while there is no set of worlds in which $\Diamond\phi$ is true, updating any particular set of worlds s with $\Diamond\phi$ will always result in a set of worlds (either s or $\emptyset$). Thus to capture modal beliefs, we propose an updated Lockean thesis: for A to believe ϕ is to assign a high credence to the result of updating A's doxastic state with ϕ. Intuitively, we can think of updating one's doxastic state with ϕ as a way of modeling becoming certain of ϕ. Given this gloss, our proposal amounts to the following. A believes ϕ iff A assigns a sufficiently high credence to the doxastic state that would result from becoming certain of ϕ.

In order to implement this, we model an agent A's doxastic state at w with two components—a set of worlds s_A^w, and a probability function Pr_A^w. s_A^w is the set of worlds compatible with what A is certain of at w. As in Lockean Belief, Pr_A^w is an agent's credence function at w. It assigns s_A^w a probability of 1. We model the result of updating A's information at w with ϕ as $s_A^w[\phi]$, where $[\cdot]$ is defined as in Update Semantics. Our proposal is that an agent believes ϕ iff she assigns a sufficiently high credence to *this* set of worlds ($s_A^w[\phi]$). And so updating a context with $B_A\phi$ ensures that it only contains worlds where A's credence function meets this condition. More precisely:

Locke Updated: $s[B_A\phi] = \{w \in s \mid Pr_A^w(s_A^w[\phi]) > t\}$.

Let's unpack this. According to Locke Updated, $B_A\phi$ updates a context with the worlds where A believes ϕ. Which worlds are these? The worlds where A has a sufficiently high credence in $s_A^w[\phi]$. Here $s_A^w[\phi]$ represents the doxastic state that A would be in, if A were to become certain of ϕ.[25]

We now show how this semantics resolves our puzzle. To do this, we first show that our semantics validates both Uncertain Belief and Uncertainty-Possibility Link, thereby validating Fallibility. We next show that our semantics validates No Contradictions.

Start with Uncertain Belief. To see that Locke Updated validates this principle, note that Locke Updated agrees with Lockean Belief when it comes to descriptive beliefs. If ϕ is descriptive, then to believe ϕ is to assign a sufficiently high credence to the ϕ worlds:

Fact 2 (Descriptive Beliefs Are Lockean). For any agent A and any descriptive sentence ϕ: $s[B_A\phi] = \{w \in s \mid Pr_A^w([\![\phi]\!]) > t\}$.

[25]Locke Updated can be complicated in various ways. For example, one could also allow the threshold to vary with the context of utterance and/or the believer's practical interests (Weatherson [2005], [2012]; Ganson [2008]; Fantl and McGrath [2009]). One could also make *believes* sensitive to a question under consideration (Yalcin [2011], [forthcoming]). We will set these complications aside, since they are not directly relevant to our puzzle.

PROOF. $B_A\phi$ holds at w iff A's credence in $s^w_A[\phi]$ exceeds t. To find $s^w_A[\phi]$, we take the set of worlds in A's doxastic state at w (s^w_A) and update this set with ϕ. When ϕ is descriptive, this is simply the result of intersecting s^w_A with the ϕ worlds ($s^w_A \cap [\![\phi]\!]$). Since every agent assigns credence 1 to the set of worlds in her doxastic state, her credence in $[\![\phi]\!]$ will equal her credence in $s^w_A[\phi]$. ⊣

From Fact 2, it's a short step to Uncertain Belief. Recall that we are conceiving of an agent's doxastic state as the set of worlds compatible with her certainties. This suggests a natural semantics for certainty ascriptions: an agent is certain of ϕ just in case her doxastic state supports ϕ. This will in turn entails that her credence in ϕ is 1. Let C_A be an operator short for *A is certain that*. This gives us the following:

CERTAINTY AS SUPPORT: $s[C_A\phi] = \{w \in s \mid s^w_A \models \phi\}$.[26]

Thus, on Locke Updated, anyone whose credence in a descriptive claim ϕ is greater than t but less than 1 will count as believing ϕ without being certain of it.

To apply our semantics, recall our larcenous friend Ari, who believes the house is empty without being certain of it. We model this by saying that, at every world in the context, Ari's doxastic state includes both worlds where the house is empty and worlds where it isn't. To simplify, suppose Ari's doxastic state consists of just our three worlds from §5: w and u, in which the house is empty, and v, in which there's someone inside. Suppose that Ari assigns a credence of .8 to $\{w, u\}$ and a credence of .2 to $\{v\}$. Finally, suppose that t, the threshold for belief, is .75. Given all of this, Locke Updated entails that (23) (repeated here as (26)) is supported:

(26) Ari believes the house is empty.

After all, the result of updating Ari's doxastic state with *the house is empty* is $\{w, u\}$, and Ari's credence in this proposition exceeds .75.

At the same time, our semantics for *certain* predicts that the following will update the context to return the empty set:

(27) Ari is certain that the house is empty.

After all, Ari's doxastic state includes a world where the house is not empty (v).

[26]In §3.1, we suggested that the truth-value of a certainty ascription depends on the context (either that of the conversation, or that of the subject of the certainty ascription). One way to implement this suggestion in the present framework would be to let the subject's doxastic alternatives vary with context, so that in high-stakes contexts s^w_A includes possibilities that are absent in low-stakes contexts. See Clarke [2013] for an account of belief along much these lines.

Let's turn to Uncertainty-Possibility Link. To see that our semantics validates this principle, note that while Locke Updated and Lockean Belief agree with respect to descriptive beliefs, they diverge with respect to beliefs about what might be the case. In particular, Locke Updated agrees with Belief as Support that modal beliefs are 'transparent': an agent believes $\Diamond\phi$ just in case her doxastic state contains a ϕ world. Summarizing:

Fact 3 (*Might* Beliefs Are Transparent). For any agent A and any descriptive sentence ϕ: $s[B_A\Diamond\phi] = \{w \in s \mid s^w_A \not\models \neg\phi\}$.

Proof. A believes $\Diamond\phi$ at w just in case she gives sufficiently high credence to $s^w_A[\Diamond\phi]$. But $s^w_A[\Diamond\phi]$ is either s^w_A or $\emptyset$, depending on whether there is a ϕ world in s^w_A. If there is, then $s^w_A[\Diamond\phi] = s^w_A$, to which A assigns credence 1. Otherwise, $s^w_A[\Diamond\phi] = \emptyset$, to which A assigns credence 0. And so A believes $\Diamond\phi$ just in case her doxastic state includes a ϕ world. ⊣

Fact 3 and Certainty as Support entail Uncertainty-Possibility Link. Given Certainty as Support, if A isn't certain that ϕ, then A's doxastic state doesn't support ϕ. And so, given Fact 3, A believes $\Diamond\neg\phi$. To illustrate, let's return to Ari. Since Ari isn't certain that the house is empty, her doxastic state contains a world where the house isn't empty (v). From Fact 3, we derive that Ari believes the house might not be empty.

Because our semantics validates both Uncertain Belief and Uncertainty-Possibility Link, it validates Fallibility as an immediate corollary. According to our semantics, anytime an agent believes ϕ without being certain that ϕ, she will also count as believing $\Diamond\neg\phi$. We summarize Facts 2 and 3 in Figure 2, which illustrates the different constraints imposed by being certain that ϕ, believing ϕ, and believing $\Diamond\phi$.

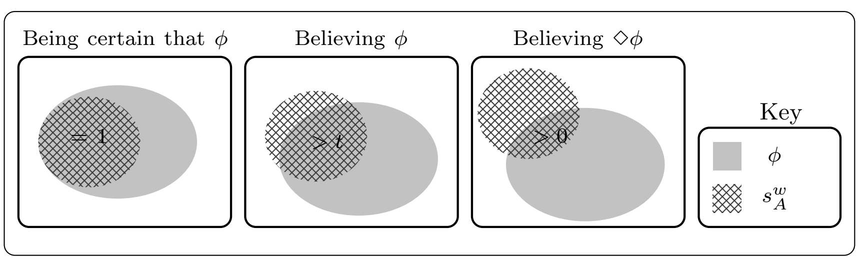

Figure 2. Locke Updated

Finally, our semantics validates No Contradictions.

Fact 4 (No Contradictions). For any context s, agent A and any descriptive sentence ϕ: $s[B_A(\phi \wedge \Diamond\neg\phi)] = \emptyset$.

PROOF. A believes $(\phi \wedge \Diamond\neg\phi)$ at w iff A assigns a sufficiently high credence to $s^w_A[\phi \wedge \Diamond\neg\phi]$. From Fact 1, we know that $s^w_A[\phi \wedge \Diamond\neg\phi] = \emptyset$. Consequently, $Pr^w_A(s^w_A[\phi \wedge \Diamond\neg\phi]) = 0$. ⊣

Applied to our example, (1) (*Ari believes the house is empty and might not be empty*) is true iff Ari assigns a sufficiently high credence to the result of updating her doxastic state with *the house is empty and might not be empty*. This proposition is found by taking her doxastic state ($\{w, u, v\}$) and updating it in two steps. First, we update it with *the house is empty*, giving us $\{w, u\}$. Next, we update this set with *the house might not be empty*, which requires checking whether $\{w, u\}$ contains at least one world where the house isn't empty. Since there's no such world in $\{w, u\}$, we get the empty set. Since coherent agents assign credence 0 to the empty set, Ari cannot coherently believe that the house is empty and might not be.

§7. Against Closure. We have proposed a new account of modal beliefs, an account that proves a solution to our puzzle.[27] We now explore a radical consequence of our solution: beliefs about what might the case furnish a new class of counterexamples to the thesis that rational belief is closed under logical implication.

Here's an off-the-shelf formulation of a multi-premise closure principle:

MULTI-PREMISE CLOSURE (MPC): If A is rational in believing a set of premises, and A competently infers some conclusion from these premises, then A's resulting belief in the conclusion is rational.[28]

[27]We claim that Locke Updated solves the puzzle, not that it is the only possible solution. To sketch just one example of an alternative approach, we could introduce a belief function $\mathcal{B}$ that, for any agent A and world w, delivers a subset of A's doxastic state at w—call this subset A's *belief state* at w. (As before, A's doxastic state at w (s^w_A) is the set of worlds compatible with A's certainties at w.) Different versions of this approach could be developed depending on how one understands the relation between A's belief state and A's doxastic state; one natural option is to take A's belief state to contain all the worlds in A's doxastic state that A takes to be most plausible—or perhaps sufficiently plausible—candidates for actuality. We could then propose that A believes ϕ iff A's belief state is a subset of the result of updating A's doxastic state with ϕ. That is, $s[B_A\phi] = \{w \in s \mid \mathcal{B}^w_A \subseteq s^w_A[\phi]\}$. This approach solves the puzzle in much the same manner as Locke Updated. It validates Uncertain Belief (since ϕ could hold throughout A's belief state without holding throughout A's doxastic state), as well as Uncertainty-Possibility Link (since if A is not certain of ϕ, then, by Certainty as Support, $s^w_A[\Diamond\neg\phi] = s^w_A$). As a consequence, it validates Fallibility. And since $s^w_A[\phi \wedge \Diamond\neg\phi] = \emptyset$, for any coherent agent A, $\mathcal{B}^w_A \not\subseteq s^w_A[\phi \wedge \Diamond\neg\phi]$; thus No Contradictions is also valid. Ultimately, the choice between Locke Updated and this alternative approach will depend—at least in part—on how plausible one finds Lockean accounts of belief in the first place. This is not an issue we will try to adjudicate here, but see the references in fn. 24 for relevant discussion. (Thanks here to Fabrizio Cariani and Johan van Benthem.)

[28]Our formulation is based on Schechter [2013].

Let ϕ_1 be the premise: *The house is empty.* Let ϕ_2 be the premise: *The house might not be empty.* Let ψ be the conclusion: *The house is empty and it might not be empty.* On our account, Ari can coherently (and presumably, rationally) believe ϕ_1 and ϕ_2, but it would be incoherent (and hence irrational) for her to believe ψ on this basis. Indeed, this is not just an idiosyncratic feature of our semantics: *any* semantics that validates both Fallibility and No Contradictions is forced to reject MPC.[29]

Of course, challenges to closure are nothing new: it's well-known that lotteries (Kyburg [1961]) and prefaces (Makinson [1965]) cause trouble for MPC. But if we're right, counterexamples to MPC are much more common than has been acknowledged. What our account suggests is that anytime you believe something without being certain of it, there's a counterexample to MPC lurking.

Here is another way in which our counterexample to MPC is stronger than others. While Bayesian theories of belief reject MPC, they accept a weaker principle (Adams [1966]; Edgington [1997]; Sturgeon [2008]). Say that an agent's uncertainty in ϕ is the difference between 1 and her credence in ϕ. So if an agent is certain that ϕ, her uncertainty in ϕ is 0. And if an agent is certain that $\neg\phi$, then her uncertainty in ϕ is 1. Bayesians accept:

Bayesian Closure (BC): If a set of premises entails some conclusion, then a rational agent's uncertainty in the conclusion cannot be greater than the sum of her uncertainty with respect to each of the premises.

According to our approach, Ari also furnishes a counterexample to BC. Since Ari's credence in ϕ_1 (*The house is empty*) is .8, her uncertainty with respect to ϕ_1 is .2. So Ari should be certain in ϕ_2 (*The house might not be empty*), which means her degree of uncertainty with respect to ϕ_2 should be 0. However, Ari should also be certain that ψ (*The house is empty and it might not be empty*) is false. So Ari's degree of uncertainty in ψ is 1, which exceeds the sum of her uncertainty in the premises ϕ_1 (.2) and ϕ_2 (0).[30]

[29] Here, we make the assumption that $\phi; \Diamond\neg\phi \models \phi \wedge \Diamond\neg\phi$. This holds on standard dynamic notions of validity such as 'update-to-test' and 'test-to-test' entailment (Veltman [1996]).

[30] Our counterexample to MPC has affinities with a recent challenge to *single-premise* closure developed independently by Bledin and Lando [forthcoming]. Bledin and Lando observe that many proponents of a non-classical semantics for epistemic modals embrace a non-classical consequence relation that validates what Yalcin [2007] refers to as "Łukasiewicz's Principle":

Łukasiewicz's Principle: $\neg\phi \models \neg\Diamond\phi$

(See, e.g., Veltman [1996]; Yalcin [2007]). Bledin and Lando point out that this principle stands in tension with single-premise closure, since it seems one can rationally believe $\neg\phi$ without being in a position to rationally believe $\neg\Diamond\phi$. For example, someone might

For those attracted to closure, it's natural to try to restrict MPC so that it does not apply to reasoning involving epistemic modals:

Restricted MPC: If A is rational in believing a set of descriptive premises, and A competently infers some descriptive conclusion from these premises, then A's resulting belief in the conclusion is rational.

Restricted MPC captures many of the intuitions that motivated closure in the first place—in particular, the idea that deduction is a rational way of extending our beliefs. On the picture that emerges, deduction is always a rational way of extending our beliefs about what *is* the case, but it's not always a rational way of extending our beliefs about what *might* be the case.

Of course, validating Restricted MPC without invalidating Uncertain Belief is no easy task. Locke Updated is not up to it. This is because Locke Updated, like Lockean Belief, embraces the idea that belief only requires meeting some threshold less than 1. Notoriously, any such Lockean approach stands in tension with even Restricted MPC.[31]

However, there are promising strategies for modifying the Lockean thesis to preserve closure. For example, Leitgeb [2014] defends a 'stability' theory of belief, according to which "Belief is determined by a proposition of resiliently or stably high subjective probability" (145). The intuitive idea here is that in order to count as believing ϕ, an agent's credence in ϕ must remain sufficiently high even upon acquiring new information (within certain limits). Leitgeb develops this idea in terms of a technical notion of 'P-Stability'. Say that a descriptive sentence ϕ is P-stable relative to a probability function Pr iff $Pr(\llbracket\phi\rrbracket)$ remains higher than $\frac{1}{2}$ even after conditionalizing on any *admissible* claim. And say that ψ is admissible (relative to ϕ, Pr) iff ψ is consistent with ϕ and Pr assigns $\llbracket\psi\rrbracket$ some non-zero

---

rationally believe that a particular lottery ticket won't win, without being in a position to rationally believe there's no possibility that it will win.

While both of our challenges involve epistemic modals, it is worth highlighting an important difference. Our challenge does not require that Łukasiewicz's Principle is valid, or even that it's rational to reason in accordance with it. All our counterexample requires is that the argument from ϕ and $\Diamond\neg\phi$ to $\phi\wedge\Diamond\neg\phi$ is valid—a point that advocates of both a classical and a non-classical consequence relation will agree on. This difference is important, because it creates difficulty for a potential response to Bledin and Lando's challenge: namely, to perform *modus tollens*, and reject any consequence relation that validates Łukasiewicz's Principle (Schulz [2010]: 388). Thus while our counterexample is strictly independent of Bledin and Lando's, we take our counterexample to complement theirs. Indeed, Locke Updated can explain why it's not always rational to form beliefs in accordance with Łukasiewicz's Principle: it's only rational to believe $\neg\Diamond\phi$ if one is certain that $\neg\phi$ is true.

[31] Note that Locke Updated does validate a restricted form of BC, according to which BC holds for descriptive premises and conclusions.

probability. That is, the set of P-stable sentences for Pr ($\mathscr{P}(Pr)$) can be characterized as follows:

P-STABILITY: $\mathscr{P}(Pr) = \{\phi \mid \forall\psi \in \mathscr{A}(\phi, Pr)\colon Pr(\llbracket\phi\rrbracket \mid \llbracket\psi\rrbracket) > \frac{1}{2}\}$, where $\psi \in \mathscr{A}(\phi, Pr)$ iff $\phi; \psi \not\models \bot$, and $Pr(\llbracket\psi\rrbracket) > 0$.

Equivalently: ϕ is P-stable just in case every world at which ϕ is true is assigned a higher probability than the union of the worlds at which ϕ is false.

As Leitgeb develops the stability theory, an agent believes ϕ just in case ϕ is entailed by some P-stable claim ψ that she believes, and hence her credence in ϕ is greater than or equal to her credence in ψ. Leitgeb shows that, on such a view, an agent's beliefs will be closed under logical implication.

As stated, Leitgeb's view is schematic. After all, a given credence function can generate multiple P-stable sentences. And for every such P-stable sentence ϕ, an agent's credence in $\llbracket\phi\rrbracket$ gives us a different candidate for the Lockean threshold. To convert Leitgeb's view into a semantics for *believes*, we can posit a contextually determined choice function f that, given the agent's set of P-stable sentences, selects one as the Lockean threshold.[32] Those attracted to this approach could modify Locke Updated as follows:

LOCKE STABILIZED: $s[B_A\phi] = \{w \in s \mid Pr^w_A(s^w_A[\phi]) \geq Pr^w_A(f(\mathscr{P}(Pr^w_A))\}$, where f selects a unique member of $\mathscr{P}(Pr^w_A)$.

To illustrate, consider (23) (*Ari believes the house is empty*). According to Locke Stabilized, this means that Ari's credence in the result of updating her doxastic state with the claim that the house is empty exceeds her credence in the f-selected P-stable claim.

While this semantics validates Restricted MPC, it does not validate an unrestricted closure principle: MPC will still fail when it comes to ϕ and $\Diamond\neg\phi$. After all, on Locke Stabilized (as on Locke Updated) an agent believes $\Diamond\phi$ iff ϕ is compatible with her doxastic state. In that case, $s^w_A[\Diamond\phi] = s^w_A$. And since $Pr^w_A(s^w_A) = 1, Pr^w_A(s^w_A|\llbracket\psi\rrbracket) = 1$ for every ψ. But $\phi \wedge \Diamond\neg\phi$ still crashes A's doxastic state, and so she does not believe this.

For our purposes, we need not commit to a stability theory of belief. Perhaps the reader prefers some other closure-preserving modification of a Lockean view. If so, we should be able to import any such modification into our semantics for *believes*, thereby preserving Restricted MPC. Or perhaps the right response to lotteries and prefaces is to abandon even restricted closure principles. For our purposes, the important point is that beliefs involving epistemic modals provide strong grounds for abandoning (unrestricted) MPC; we leave it as an open question whether our semantics

[32] See Leitgeb [2013] for the view that f selects the strongest P-stable proposition.

for *believes* should validate Restricted MPC (and, if so, how this validation is best achieved).[33]

§8. Objections.

8.1. First Objection: Order Effects. While epistemic contradictions are semantically inconsistent in Update Semantics, reversing the order of the conjuncts restores consistency:

(28) $\Diamond\neg\phi \wedge \phi$

To see that these reversed epistemic contradictions are consistent in Update Semantics, consider an example such as:

(29) ? The house might not be empty and it is empty.

And consider again our context s containing three worlds—w and u, in which the house is empty, and v, in which it isn't empty. Updating with the first conjunct of (29) (*The house might not be empty*) leaves the context the same, while updating with the second conjunct shrinks it down to $\{w, u\}$. So $s[(29)] \neq \emptyset$.

As a result, our solution to the puzzle does not predict that it's incoherent to believe reversed epistemic contradictions. That is, our semantics does not predict that instances of the following are infelicitous:

(30) $B_A(\Diamond\neg\phi \wedge \phi)$

To see this, suppose (as before) that Ari's doxastic state is $\{w, u, v\}$. Then:

(31) ? Ari believes that the house might not be empty and (it) is empty.

is predicted to narrow down the context to those worlds in which Ari's credence in $\{w, u\}$ exceeds the Lockean threshold.

Some may regard this as a problem for our approach. After all, (31) sounds fairly odd. Thus some might insist that any adequate solution to our puzzle will explain not just the incoherence of believing epistemic contradictions, but also the incoherence of believing reversed epistemic contradictions.

In response, we should first note that there is a delicate and—to our knowledge—currently unresolved question as to what exactly the data are. Historically, many dynamic semanticists have regarded it as a datum that discourse (32) is coherent, or at least less degraded than (33):

[33]Here the non-Lockean alternative developed in fn. 27 makes for an interesting comparison. On this alternative semantics, A believes a descriptive claim ϕ iff ϕ holds throughout all the worlds in A's belief state. Thus this alternative approach agrees with Locke Stabilized in validating Restricted MPC. And like Locke Stabilized, it still invalidates Unrestricted MPC, since it still predicts failures of closure when it comes to ϕ and $\Diamond\neg\phi$.

(32) It might be raining. It isn't raining.

(33) ? It isn't raining. It might be raining.[34]

In a similar vein, Sorensen [2009] and Dorr and Hawthorne [2013] claim that reversing the order of the conjuncts of an embedded epistemic contradiction tends to make the sentence more acceptable. Others have questioned these judgments.[35] Until this empirical issue is investigated more fully, it's not obvious whether the fact that our theory predicts that (30) is consistent should be regarded as a vice or a virtue.

But suppose we set aside this question about the data and assume, at least for the sake of argument, that instances of (30) are typically judged infelicitous. Is there any way of modifying our semantics to predict this? Note that while reversed epistemic contradictions are consistent, they do display a somewhat unusual property. Say that a sentence is *incohesive* if there is no non-absurd context that supports it:

Cohesion: A sentence ϕ is cohesive iff $\exists s \neq \emptyset$: $s \models \phi$; otherwise ϕ is incohesive.[36]

It is easy to show that reversed epistemic contradictions are incohesive:

Fact 5 (Reversed Epistemic Contradictions are Incohesive). For any descriptive sentence ϕ and any non-absurd context s: $s \not\models \Diamond\neg\phi \wedge \phi$.

Proof. Either s contains a $\neg\phi$ world or it does not. If it does, $s[\Diamond\neg\phi \wedge \phi] \neq s$, since updating with the second conjunct will remove the $\neg\phi$ world. And if s does not contain a $\neg\phi$ world, then $s[\Diamond\neg\phi \wedge \phi] \neq s$, since the first conjunct will crash the context. Either way, $s \not\models \Diamond\neg\phi \wedge \phi$. ⊣

As a number of authors have noted, our reaction to (29) suggests that there is something defective—or at least odd—about asserting incohesive sentences.[37] And our reaction to (31) suggests that this defect also extends to belief reports that embed incohesive sentences. The question is now whether there is any way of incorporating this observation into our semantics.

Two strategies suggest themselves. Taking a cue from Klinedinst and Rothschild [2015] and Yalcin [2015], the first option is to modify the update function in Update Semantics to predict that sentences containing

[34]See, for example, Groenendijk and Stokhof [1991b]; Veltman [1996]; Gillies [2001]; Fintel and Gillies [2007].

[35]See, for example, Yalcin [2015].

[36]This property is often referred to as 'coherence' (Groenendijk, Stokhof, and Veltman [1996]; Willer [2013]). We follow Gillies [2004] in using the label 'cohesiveness' so as to avoid confusion with the intuitive notion of coherence deployed in our earlier principles.

[37]Or at least this is odd when no new information is received between the utterance of the first and second conjunct (Groenendijk, Stokhof, and Veltman [1996]; Gillies [2004]; Willer [2013]).

incohesive constituents always crash the context.[38] A second option is to modify not the update function but the semantics for *believes*. As an anonymous referee helpfully observes, one way to do so would be to hold that A believes ϕ iff there's some subset of their doxastic state that meets two conditions: (i) A assigns it a sufficiently high probability, (ii) it supports ϕ. That is:

LOCKE SUPPORTED: $s[B_A\phi] = \{w \in s \mid \exists s' \subseteq s^w_A : Pr^w_A(s') > t \ \&\ s' \models \phi\}$

Locke Supported validates all of our principles from §3. It agrees with both Lockean Belief and Locke Updated that believing a descriptive claim ϕ only requires assigning a sufficiently high probability to the set of ϕ worlds; it thus validates Uncertain Belief. And it agrees with Locke Updated that modal beliefs are transparent—i.e., A believes $\Diamond\neg\phi$ iff there's at least one $\neg\phi$ world in A's doxastic state. It thus validates Uncertainty-Possibility Link. As a consequence of these two facts, Locke Supported preserves Fallibility. Finally, since no set of worlds supports an epistemic contradiction, Locke Supported also predicts that no agent believes an epistemic contradiction. At the same time, Locke Supported avoids order effects. After all, since reversed epistemic contradictions are incohesive, the second condition of Locke Supported ensures that no coherent agent believes a reversed epistemic contradiction.

Is one of these strategies—modifying the update function, or modifying the semantics for *believes*—preferrable to the other? For the purposes of this paper, we will refrain from taking a stand on this question. However, let us briefly raise one pertinent consideration. Note that Locke Supported and Locke Updated differ not only with respect to reversed epistemic contradictions, but also with respect to beliefs about what *must* be the case. After all, in Update Semantics, any state that supports ϕ will also support $\neg\Diamond\neg\phi$, and hence, assuming *must* ($\Box$) is the dual of $\Diamond$, will also support $\Box\phi$. So Locke Supported predicts that anyone who believes ϕ also believes $\Box\phi$. By contrast, Locke Updated predicts that an agent only believes $\Box\phi$ if they are certain of ϕ. To test which of these predictions is correct, consider variants of our earlier examples (4) and (7) involving necessity beliefs:

(34) ? Ari believes the house must be empty. But she realizes that it might not be.

[38] Klinedinst and Rothschild [2015] and Yalcin [2015] locate the trouble in a slightly weaker property of reversed epistemic contradictions, specifically, their *non-idempotence*, where a sentence is idempotent iff updating any context with it once delivers the same result as updating the context with it twice—i.e., $\forall s : s[\phi] \models \phi$. (Note that any incohesive sentence is non-idempotent, but not *vice versa*.) See Yalcin [2015]: 503-504 for one way of modifying the update function to rule out sentences that are defined in terms of non-idempotent updates.

(35) ? Ari believes there's no possibility the house isn't empty. But she isn't certain the house is empty.

To our ears such discourses sound odd—worse, at any rate, than their necessity modal-free counterparts. Insofar as others share these intuitions, this provides reason to resist replacing Locke Updated with Locke Supported,[39] and hence to explain the oddity of (31) via the first strategy.[40]

8.2. Second Objection: Questioning Fallibility. According to Locke Updated, believing a conjunction is not equivalent to believing each conjunct individually. In particular, while Locke Updated predicts that $B_A(\phi \wedge \Diamond\neg\phi)$ crashes, it doesn't predict that (36) crashes:

(36) $B_A\phi$. $B_A\Diamond\neg\phi$.

Many find discourses like this at least somewhat odd. Consider, for instance:

(37) ? Ari believes the house is empty. She also believes it might not be empty.

Should we revise our semantics to make (37) inconsistent after all?

Here too there is a question about the data. While we agree that (37) is a bit peculiar, it doesn't sound to us *as* bad as (1). Informal polling suggests others agree: roughly half of those we've surveyed judge (37) to be marked, while the other half judged it to be fine. By contrast, the vast majority of respondents deem (1) infelicitous. This provides reason to resist revising

[39]See Bledin and Lando [forthcoming] for further arguments (mentioned in fn. 30) against the view that believing ϕ entails believing $\Box\phi$.

[40]These two strategies are not the only options. Another route is to modify the semantics for conjunction. One simple modification would be to make $\wedge$ perform a second, reverse sequential upate—that is: $s[\phi \wedge \psi] = s[\phi][\psi] \cap s[\psi][\phi]$. This ensures that reversed epistemic contradictions are inconsistent, because computing $s[\Diamond\neg\phi \wedge \phi]$ would require computing $s[\phi][\Diamond\neg\phi]$, the usual epistemic contradiction. However, we worry that this strategy may prove too extreme, since a commutative semantics for conjunction along these lines would rule out order effects involving other phenomena that have called for a dynamic treatment—for example, presupposition and anaphora.

A more nuanced modification of the semantics for conjunction may fare better. For example, we could exploit Fact 5 by letting $\wedge$ return the maximal subset of the context that supports the standard dynamic conjunction. That is: $s[\phi \wedge \psi] = \bigcup\{s' \subseteq s \mid s'[\phi][\psi] = s'\}$. This semantics is not commutative; it thus allows for order effects involving presupposition and anaphora. At the same time, reversed epistemic contradictions are still guaranteed to crash every context, since no context supports them. However, this semantics has some surprising consequences. For example, it predicts that $\Box\phi$ is not equivalent to $\Box\phi \wedge \Box\phi$. (To see this, consider again our context $s = \{w, u, v\}$. Note that $s[\Box\phi \wedge \Box\phi] = \{w, u\}$, since this is the maximal subset of s that supports $\Box\phi$. But $s[\Box\phi] = \emptyset$, since s contains v, which is a $\neg\phi$ world.) We leave it to future work to explore whether this consequence is acceptable, and—if not—whether some other modification of the semantics for conjunction proves more successful.

our semantics to predict that instances of (36) crash: an adequate solution to the puzzle will capture the fact that (1) sounds worse than (37).

In addition to our scruples about the infelicity of (37), there are principled reasons to resist revising our semantics to predict that instances of (36) crash. Any such revision would require giving up Fallibility. But, as we have seen (§3.1), there are independent arguments for Fallibility. To review: there was the argument from the felicity of CBAs (e.g., (3)); there was the argument from the felicity of variants of (37) involving *realize* and *recognize* (e.g., (4)); finally, there was the theoretical argument, which derived Fallibility from the conjunction of Uncertain Belief and Uncertainty-Possibility Link, both of which were supported by independent data.

For ease of reference, (38) collects much of the relevant data in one place, showcasing these three grades of modal infelicity:

(38)
a. # A believes $(\phi \wedge \Diamond\neg\phi)$.
b. ? A believes ϕ. A also believes $\Diamond\neg\phi$.
c. ✓ A believes ϕ. But A realizes $\Diamond\neg\phi$.
d. ✓ I believe ϕ. But $\Diamond\neg\phi$.

None of the views that we've considered predicts all of these judgments. However, Locke Updated comes the closest. On the one hand, views that reject Fallibility (e.g., the belief-based version of the classical semantics and Belief as Support) incorrectly predict that all of these sentences are infelicitous. On the other hand, the knowledge-based version of the classical semantics does not predict that any of these sentences are infelicitous. Only Locke Updated predicts both the infelicity of (38a) and the felicity of (38c)-(38d).

Still, our initial question remains: what explains why (37/38b) is at least somewhat odd? We are not sure, but one hypothesis is that optional modal subordination is responsible.[41] A modal is subordinated when it is evaluated relative to a range of possibilities controlled by previous discourse. We propose that when we evaluate (37/38b), we tend to access a reading that subordinates the modal *might* to the complement of the previous belief report (*The house is empty*). On this reading, the modal is evaluated relative to a set of worlds where the house is empty; consequently, the discourse crashes. While this reading is available, it is not mandatory; there's another reading of (37/38b) on which the modal is not subordinated. On this reading, the discourse does not crash. We suspect the availability of this unsubordinated reading explains why (37) is less odd than (1).

[41]For discussion of modal subordination, see Roberts [1989]; Kibble [1995]; Rooij [2005]. For evidence that *might* gives rise to optional modal subordination, see Klecha [2012].

This explanation leaves some questions unanswered. First, why do some speakers prefer the subordinated reading over the unsubordinated reading? Second, why does (4/38c) sound better than (37/38b)? We do not at present have fully developed answers to these questions. We suspect that when a modal has two readings—one subordinated, the other unsubordinated—the extent to which one reading will be preferred over the other will be influenced by a variety of factors. Perhaps for some speakers subordinated readings are the default—the readings they tend to latch onto in the absence of cues to the contrary. But this default can be overridden. Tellingly, (4/38c) contains a contrast marker (*but*); what's more, this discourse sounds best when the contrast marker is given prosodic focus. We suspect that this may override the default, biasing speakers towards the unsubordinated reading.[42] However, we will leave to future research the project of developing a rigorous account of the factors that influence the extent to which the different readings are preferred. What's important for our purposes is that there are compelling reasons to retain Fallibility. This suggests that the right response to (37/38b) is not to revise Locke Updated, but rather to explain the oddity of these sentences via some pragmatic mechanism.

8.3. Third Objection: *Might* Beliefs vs. *Might* Certainties. According to our view, an agent believes $\Diamond\phi$ iff ϕ is compatible with her certainties. Given Certainty as Support, an agent is certain of $\Diamond\phi$ iff ϕ is compatible with her certainties. Thus our proposal collapses believing ϕ is possible and being certain that ϕ is possible:

FACT 6 (Collapse). For any context s, agent A and any descriptive sentence ϕ: $s[B_A\Diamond\phi] = s[C_A\Diamond\phi]$.

Proof: From Fact 2, $s[B_A\Diamond\phi] = \{w \in s \mid s^w_A \not\models \neg\phi\}$. By Certainty as Support and Update Semantics, $s[C_A\Diamond\phi] = \{w \in s \mid s^w_A \not\models \neg\phi\}$. So $s[B_A\Diamond\phi] = s[C_A\Diamond\phi]$.

This seems counterintuitive. Consider, for example, DeRose [1991]'s cancer case, in which Jane's husband John has undergone a test to determine whether he has cancer. A negative result will mean that John definitely does not have cancer. A positive result does not necessarily mean that John does have cancer; rather, it means that further tests have to be run. It seems natural to describe Jane's credal state as follows:

(39) Jane believes John might have cancer. But she isn't certain he might have cancer.

[42] It's well known that contrast—and coherence relations more generally—influence how we resolve ambiguous sentences (Hobbs [1985]; Kehler [2002]). For discussion of how coherence relations influence the resolution of epistemic modals in particular, see Asher and McCready [2007].

One possible response is to claim that the modal in the second conjunct is not epistemic; rather, it quantifiers over physical or metaphysical possibilities (Stephenson [2007b]: 50). However, some might regard this as *ad hoc*. At the very least, it doesn't seem obvious that the occurrence of *might* in the second conjunct is non-epistemic.

Perhaps a better response is to concede the counterexample and amend our semantics for certainty ascriptions. Earlier we raised the possibility that there is a stability constraint on belief. We could likewise impose a stability constraint on certainty. According to this constraint, in order for A to be certain that ϕ, A's doxastic state must support ϕ even once it's been updated with any admissible claim. As before, ψ is admissible iff ψ is consistent with ϕ and assigned some non-zero credence by A. That is:

Certainty Stabilized: $s[C_A\phi] = \{w \in s \mid \forall\psi \in \mathscr{A}(\phi, Pr^w_A)\colon\ s^w_A[\psi] \models \phi\}$, where $\psi \in \mathscr{A}(\phi, Pr^w_A)$ iff $\phi;\psi \not\models \bot$ and $Pr^w_A([\![\psi]\!]) > 0$.[43]

To see how this solves the present difficulty, let ψ be the claim that the test results are negative. This is consistent with the claim that John might have cancer; Jane also assigns this claim some non-zero credence. But if we update Jane's doxastic state with ψ, the resulting set contains no worlds in which John has cancer. And so the resulting set will fail the test imposed by the sentence: *John might have cancer.* Thus Certainty Stabilized predicts that the second sentence in (39) (*She isn't certain he might have cancer*) is true, as desired.

Certainty Stabilized is a conservative extension of Certainty as Support. That is, the two semantics make the same predictions whenever ϕ is not a possibility claim. To see why, we need to introduce some further terminology. A sentence ϕ is *persistent* when any context s that supports ϕ will continue to support ϕ once s is updated with more information (Veltman [1996]: 3):

Persistence: ϕ is persistent iff $\forall s\ \forall\psi$: if $s \models \phi$, then $s[\psi] \models \phi$.

In Update Semantics, any descriptive claim is persistent. After all, s supports some descriptive claim ϕ just in case ϕ holds at every world in s. Whenever this obtains, ϕ also holds at any subset of s. By contrast, $\Diamond\phi$ is not persistent (Veltman [1996]). After all, if $s \models \Diamond\phi$, then updating s with ϕ will not yield $\emptyset$. But it may well be that updating certain subsets of s with ϕ will yield $\emptyset$. For example, when $s \models \Diamond\phi$ we will not in general have that $s[\neg\phi] \models \Diamond\phi$.

[43]This corresponds to the claim that A is certain of ϕ iff ϕ is P^1*-stable*, where ϕ is P^1-stable iff, for any sentence ψ that is both consistent with ϕ and assigned some non-zero by A, $Pr(s^w_A[\phi] \mid [\![\psi]\!] = 1)$. For further discussion of different levels of P-stability, see Leitgeb [forthcoming] (appendix B).

It turns out that whenever ϕ is persistent, Certainty Stabilized agrees with Certainty as Support. That is, assuming the semantics for $C_A\phi$ is provided by Certainty Stabilized, the following holds:

FACT 7 (Persistent Certainties are Stable). For any agent A and any persistent sentence ϕ: $s[C_A\phi] = \{w \in s \mid s^w_A \models \phi\}$.

PROOF. It suffices to show that if ϕ is persistent, then $s^w_A \models \phi$ iff $\forall\psi \in \mathscr{A}(\phi, Pr^w_A) :\ s^w_A[\psi] \models \phi$. So suppose $s^w_A \models \phi$. Then since ϕ is persistent, $\forall\psi \in \mathscr{A}(\phi, Pr^w_A) :\ s^w_A[\psi] \models \phi$. Similarly, suppose that $\forall\psi \in \mathscr{A}(\phi, Pr^w_A)$: $s^w_A[\psi] \models \phi$. Then since $\top \in \mathscr{A}(\phi, Pr^w_A), s^w_A \models \phi$. ⊣

This last fact leads to an interesting prediction. While $\Diamond\phi$ is not persistent in Update Semantics, $\Box\phi$ is.[44] After all, $s \models \Box\phi$ just in case $s \models \phi$. If this holds, then by the persistence of ϕ we know that for any $\psi : s[\psi] \models \phi$, and hence that $s[\psi] \models \Box\phi$. Thus while Certainty Stabilized allows for a distinction between belief and certainty about what might be the case, it doesn't allow for any distinction between belief and certainty about what *must* be the case. Interestingly, this prediction appears to be borne out by the data. Consider the variant of (39) that replaces *might* with *must*:

(40) ?? Jane believes John must have cancer. But she isn't certain he must have cancer.

We find it much harder to access a true reading of (40) than (39). This is a surprising observation—one that Certainty Stabilized elegantly explains.

§9. Conclusion. Recent work on the semantics of epistemic modals has explored what sort of mental state is involved in beliefs about epistemic possibility: what is it to believe that something *might* be the case? In this paper, we've tried to make progress on this question. We began by identifying two principles that should constrain any account of what's involved in believing that something might be the case. Taken together, these principles form a puzzle, since they are *prima facie* difficult to reconcile. We went on to resolve this puzzle by offering a new semantics for *believes* that integrates a Bayesian approach to belief with a dynamic semantics for modals.

While we have focused on belief, our approach extends to other attitudes. Say that an attitude verb V is *credal* iff $V_A\phi$ entails that A assigns some non-zero credence to ϕ. Many—perhaps most—attitude verbs are credal: one cannot suspect, regret, fear, or hope the house is empty if one is certain the house isn't empty. A natural generalization of our semantics for *believes* holds that for any credal attitude verb V, $V_A\phi$ entails that A assigns some

[44]Here we again assume that $\Box$ is the dual of $\Diamond$, and so $s[\Box\phi] = \{w \in s \mid s \models \phi\}$.

credence to the result of updating A's doxastic state with ϕ (i.e., $Pr^w_A(s^w_A[\phi])$ > 0).

This proposal provides a general explanation of the oddity of epistemic contradictions under credal attitude verbs:

(41) # Ari suspects/regrets/fears/hopes that [the house is empty and might not be].

Since A will —if coherent—always have credence 0 in the result of updating her doxastic state with an epistemic contradiction, this proposal explains the incoherence of adopting any credal attitude towards an epistemic contradiction.[45] Thus the account of belief developed in this paper has repercussions for our understanding of a broader class of attitudes.

References

Ernest Adams [1966], *Probability and the logic of conditionals*, ***Aspects of inductive logic*** (Hintikka and Suppes, editors), North-Holland, Amsterdam, pp. 165–316.

Pranav Anand and Valentine Hacquard [2013], *Epistemics and attitudes*, ***Semantics and Pragmatics***, vol. 6, pp. 1–59.

D.M. Armstrong [1973], ***Belief, truth, and knowledge***, Cambridge University Press, London.

Nicholas Asher and Eric McCready [2007], Were*,* would*,* might *and a compositional account of counterfactuals*, ***Journal of Semantics***, vol. 24, pp. 93–129.

A.J. Ayer [1936], ***Language, truth, and logic***, Dover, New York.

Bob Beddor [2016], ***Reduction in epistemology***, PhD thesis, Rutgers University, New Brunswick, NJ.

Alexander Bird [2007], *Justified judging*, ***Philosophy and Phenomenological Research***, vol. 74, pp. 81–110.

Justin Bledin and Tamar Lando [forthcoming], *Closure and epistemic modals*, ***Philosophy and Phenomenological Research***.

David Christensen [2005], ***Putting logic in its place***, Oxford University Press.

Roger Clarke [2013], *Belief is credence one (in context)*, ***Philosophers' Imprint***, vol. 13, no. 11, pp. 1–18.

Claudio de Almeida [2001], *What Moore's paradox is about*, ***Philosophy and Phenomenological Research***, vol. 62, no. 1, pp. 33–58.

Keith DeRose [1991], *Epistemic possibilities*, ***Philosophical Review***, vol. 100, no. 4, pp. 581–605.

Dylan Dodd [2011], *Against fallibilism*, ***Australasian Journal of Philosophy***, vol. 89, no. 4, pp. 665–685.

Dylan Dodd [forthcoming], *Belief and certainty*, ***Synthese***.

Cian Dorr and John Hawthorne [2013], *Embedding epistemic modals*, ***Mind***, vol. 488, no. 122, pp. 867–913.

[45] What about non-credal attitude verbs? One option is to generalize the semantics for *supposes* in fn.22. For any non-credal attitude verb N, let N^w_A be the set of worlds compatible with what A N's at w. Then: $s[N_A\phi] = \{w \in s \mid N^w_A \models \phi\}$. This predicts that one cannot coherently take any non-credal attitude towards an epistemic contradiction, since no set of worlds supports an epistemic contradiction.

Kevin Dorst [forthcoming], *Lockeans maximize expected accuracy*, ***Mind***.

J.L. Dowell [2011], *A flexible contextualist account of epistemic modals*, ***Philosophers' Imprint***, vol. 11, no. 14, pp. 1–25.

Dorothy Edgington [1997], *Vagueness by degrees*, ***Vagueness: A reader*** (Keefe and Smith, editors), MIT Press, Cambridge, MA.

Andy Egan [2007], *Epistemic modals, relativism, and assertion*, ***Philosophical Studies***, vol. 133, no. 1, pp. 1–22.

Andy Egan, John Hawthorne, and Brian Weatherson [2005], *Epistemic modals in context*, ***Contextualism in philosophy*** (Preyer and Peter, editors), Oxford University Press, Oxford.

Andy Egan and Brian Weatherson [2011], *Epistemic modals and epistemic modality*, ***Epistemic modality*** (Egan and Weatherson, editors), Oxford University Press, Oxford.

Jeremy Fantl and Matthew McGrath [2009], ***Knowledge in an uncertain world***, Oxford University Press, Oxford.

Richard Foley [1993], ***Working without a net***, Oxford University Press.

Richard Foley [2009], *Belief, degrees of belief, and the lockean thesis*, ***Degrees of belief*** (Huber and Schmidt-Petri, editors), Springer, pp. 37–47.

Dorit Ganson [2008], *Evidentialism and pragmatic constraints on outright belief*, ***Philosophical Studies***, vol. 139, pp. 441–458.

Anthony Gillies [2001], *A new solution to moore's paradox*, ***Philosophical Studies***, vol. 105, no. 3, pp. 237–250.

Anthony Gillies [2004], *Epistemic conditionals and conditional epistemics*, ***Noûs***, vol. 38, no. 4, pp. 585–616.

Michael Green and John N. Williams (editors) [2007], ***New essays on moore's paradox***, Oxford University Press, New York.

Jeroen Groenendijk and Martin Stokhof [1990], *Dynamic montague grammar*, ***Papers from the second symposium on logic and language*** (Budapest) (Laszlo Kalman and Laszlo Polos, editors), Akademiai Kiado, pp. 3–48.

Jeroen Groenendijk and Martin Stokhof [1991a], *Dynamic predicate logic*, ***Linguistics and Philosophy***, vol. 14, pp. 39–100.

Jeroen Groenendijk and Martin Stokhof [1991b], *Two theories of dynamic semantics*, ***Lecture Notes in Computer Science***, vol. 478, pp. 55–64.

Jeroen Groenendijk, Martin Stokhof, and Frank Veltman [1996], *Coreference and modality*, ***The handbook of contemporary semantic theory*** (Shalom Lappin, editor), Blackwell, Cambridge, MA, pp. 179–214.

Ian Hacking [1967], *Possibility*, ***Philosophical Review***, vol. 76, pp. 143–168.

John Hawthorne [2007], *Eavesdroppers and epistemic modals*, ***Philosophical Issues***, vol. 17, pp. 92–101.

John Hawthorne [2012], *Knowledge and epistemic necessity*, ***Philosophical Studies***, vol. 158, no. 3, pp. 493–501.

John Hawthorne, Daniel Rothschild, and Levi Spectre [2016], *Belief is Weak*, ***Philosophical Studies***, vol. 173, pp. 1393–1404.

Irene Heim [1982], *The semantics of definite and indefinite noun phrases*, ***Ph.D. thesis***, University of Massachusetts, Amherst, MA.

Irene Heim [1983], *On the projection problem for presuppositions*, ***Wccfl***, vol. 2, pp. 114–125.

Irene Heim [1992], *Presupposition projection and the semantics of attitude verbs*, ***Journal of Semantics***, vol. 9, no. 3, pp. 183–221.

Jaakko Hintikka [1962], ***Knowledge and belief***, Cornell University Press.

Jerry R. Hobbs [1985], *On the coherence and structure of discourse*, ***Technical Report CSLI-85-37***, Stanford University, Center for the Study of Language and Information.

Wesley Holliday and Thomas Icard [2010], *Moorean phenomena in epistemic logic*, (Beklemishev, Goranko, and Shehtman, editors), vol. 8, College Publications, London, pp. 178–199.

Michael Huemer [2007], *Moore's paradox and the norm of belief*, ***Themes from g.e. moore: New essays in epistemology and ethics*** (Nuccetelli and Seay, editors), vol. 74, Clarendon Press, Oxford, pp. 142–157.

Lauri Karttunen [1974], *Presuppositions and linguistic context*, ***Theoretical Linguistics***, vol. 1, pp. 181–194.

Andrew Kehler [2002], ***Coherence, reference, and the theory of grammar***, CSLI Publications.

Chris Kennedy [2007], *The semantics of relative and absolute gradable adjectives*, ***Linguistics and Philosophy***, vol. 30, pp. 1–45.

Rodger Kibble [1995], *Dynamics of epistemic modality and anaphora*, ***Proceedings of the International Workshop on Computational Semantics, ITK***, pp. 121–130.

Peter Klecha [2012], *Positive and conditional semantics for gradable modals*, ***Proceedings of sinn und bedeutung 16*** (Guevara et al, editor), vol. 2, MIT Working Papers in Linguistics, pp. 363–376.

Nathan Klinedinst and Daniel Rothschild [2015], *Quantified epistemic modality*, Manuscript.

Angelika Kratzer [1981], *The notional category of modality*, ***Words, worlds, and contexts: New approaches in word semantics*** (Eikmeyer and Rieser, editors), W. de Gruyter, Berlin.

Angelika Kratzer [2012], ***Modals and conditionals***, Oxford University Press, Oxford.

Henry Kyburg [1961], ***Probability and the logic of rational belief***, Wesleyan University Press, Middletown, CT.

Peter Lasersohn [1999], *Pragmatic halos*, ***Language***, vol. 75, pp. 522–551.

Hannes Leitgeb [2013], *Reducing belief simpliciter to degrees of belief*, ***Annals of Pure and Applied Logic***, vol. 164, pp. 1338–1389.

Hannes Leitgeb [2014], *The stability theory of belief*, ***Philosophical Review***, vol. 123, no. 3, pp. 131–171.

Hannes Leitgeb [forthcoming], ***The stability of belief***, Oxford University Press, New York.

David Lewis [1979], *Scorekeeping in a language game*, ***Journal of Philosophical Logic***, vol. 8, no. 1, pp. 339–359.

David Lewis [1996], *Elusive knowledge*, ***Australasian Journal of Philosophy***, vol. 74, no. 4, pp. 549–567.

John MacFarlane [2011], *Epistemic modals are assessment sensitive*, ***Epistemic modality*** (Egan and Weatherson, editors), Oxford University Press, Oxford.

David Makinson [1965], *The paradox of the preface*, ***Analysis***, vol. 25, no. 6, pp. 205–207.

Aidan McGlynn [2013], *Believing things unknown*, ***Noûs***, vol. 47, no. 2, pp. 385–407.

Aidan McGlynn [2014], ***Knowledge first?***, Palgrave McMillian, New York.

G.E. Moore [1959], *Certainty*, ***Philosophical papers***, Allen and Unwin, London.

Sarah Moss [2015], *On the semantics and pragmatics of epistemic vocabulary*, ***Semantics and Pragmatics***, vol. 8, no. 5, pp. 1–81.

Richard Pettigrew [2015], *Accuracy and the credence-belief connection*, ***Philosophers' Imprint***, vol. 15, no. 16, pp. 1–20.

Colin Radford [1966], *Knowledge: By examples*, ***Analysis***, vol. 27, pp. 1–11.

Craige Roberts [1989], *Modal subordination and pronominal anaphora in discourse*, ***Linguistics and Philosophy***, vol. 12, no. 6, pp. 683–721.

Bertrand Russell [1912], ***The problems of philosophy***, Henry Holt and Company, New York.

Patrick Rysiew [2001], *The context-sensitivity of knowledge attributions*, ***Noûs***, vol. 35, no. 4, pp. 477–514.

Joshua Schechter [2013], *Rational self-doubt and the failure of closure*, ***Philosophical Studies***, vol. 163, no. 2, pp. 428–452.

Moritz Schulz [2010], *Epistemic modals and informational consequence*, ***Synthese***, vol. 174, pp. 385–395.

Roy Sorensen [2009], *Meta-agnosticism: Higher order epistemic possibility*, ***Mind***, vol. 118, no. 471, pp. 777–784.

Robert Stalnaker [1973], *Presuppositions*, ***Journal of Philosophical Logic***, vol. 2, pp. 447–457.

Jason Stanley [2005], *Fallibilism and concessive knowledge attributions*, ***Analysis***, vol. 65, no. 2, pp. 126–131.

Jason Stanley [2008], *Knowledge and certainty*, ***Philosophical Issues***, vol. 18, no. 1, pp. 35–57.

Tamina Stephenson [2007a], *Judge dependence, epistemic modals, and predicates of personal taste*, ***Linguistics and Philosophy***, vol. 30, no. 4, pp. 487–525.

Tamina Stephenson [2007b], ***Towards a theory of subjective meaning***, Ph.D. Thesis, MIT.

Scott Sturgeon [2008], *Reason and the grain of belief*, ***Noûs***, vol. 42, no. 1, pp. 359–396.

Jonathan Sutton [2005], *Stick to what you know*, ***Noûs***, vol. 39, no. 3, pp. 359–396.

Jonathan Sutton [2007], ***Without justification***, MIT Press, Cambridge, MA.

Peter Unger [1975], ***Ignorance***, Clarendon Press, Oxford.

Robert van Rooij [2005], *A modal analysis of presupposition and modal subordination*, ***Journal of Semantics***, pp. 281–305.

Frank Veltman [1996], *Defaults in update semantics*, ***Journal of Philosophical Logic***, vol. 25, no. 3, pp. 221–261.

Kai von Fintel and Anthony Gillies [2007], *An opinionated guide to epistemic modality*, ***Oxford studies in epistemology*** (Gendler and Hawthorne, editors), vol. 2, Oxford University Press, Oxford, pp. 32–62.

Kai von Fintel and Anthony Gillies [2010], Must*... stay... strong!*, ***Natural Language Semantics***, vol. 18, no. 4, pp. 351–383.

Brian Weatherson [2005], *Can we do without pragmatic encroachment?*, ***Philosophical Perspectives***, vol. 19, pp. 417–443.

Brian Weatherson [2012], *Knowledge, bets, and interests*, ***New essays on knowledge ascriptions*** (Brown and Gerken, editors), Oxford University Press, New York, pp. 75–103.

Malte Willer [2013], *Dynamics of epistemic modality*, ***Philosophical Review***, vol. 122, no. 1, pp. 44–92.

John N Williams [1994], *Moorean absurdity and the intentional 'structure' of assertion*, ***Analysis***, vol. 54, no. 3, pp. 24–22.

Timothy Williamson [2000], ***Knowledge and its limits***, Oxford University Press, Oxford.

Alex Worsnip [2015], *Possibly false knowledge*, ***Journal of Philosophy***, vol. 112, pp. 225–246.

Seth Yalcin [2007], *Epistemic modals*, ***Mind***, vol. 116, no. 464, pp. 983–1026.

Seth Yalcin [2011], *Nonfactualism about epistemic modality*, ***Epistemic modality*** (Egan and Weatherson, editors), Oxford University Press, Oxford.

Seth Yalcin [2012a], *Context probabilism*, ***Proceedings of the 18th Amsterdam Colloquium***, pp. 12–21.

Seth Yalcin [2012b], *A counterexample to modus tollens*, ***Journal of Philosophical Logic***, vol. 41, no. 6, pp. 1001–1024.

Seth Yalcin [2015], *Epistemic modality* De Re, ***Ergo***, vol. 2, no. 19, pp. 475–527.

Seth Yalcin [forthcoming], *Belief as question-sensitive*, ***Philosophy and Phenomenological Research***.

Henk Zeevat [1992], *Presupposition and accommodation in update semantics*, ***Journal of Semantics***, vol. 9, no. 2, pp. 379–412.