On the semantics and pragmatics of epistemic vocabulary*

Sarah Moss
University of Michigan

Abstract This paper motivates and develops a novel semantics for several epistemic expressions, including possibility and necessity modals and indicative conditionals. The semantics I defend constitutes an alternative to standard truth conditional theories, as it assigns sets of probability measures rather than sets of worlds as sentential semantic values. I argue that what my theory lacks in conservatism, it makes up for in strength — namely, the theory accounts for a host of distinctive and suggestive linguistic data collected and explored in this paper.

There has been much recent debate over the correct semantics for epistemic vocabulary — that is, expressions like the sentential operators in sentences such as:

(1)   John might be in his office.

(2)   John must be in his office.

(3)   John is probably in his office.

(4)   If John is in the building, he is in his office.

This paper explores a rich source of data for theories of this vocabulary. The debate over the viability of standard truth conditional theories has called attention to the distinctive behavior of epistemic vocabulary in eavesdropping

---

judgments, indicative suppositions, and statements of disagreement and retraction. But extant accounts are not sufficiently sensitive to distinctive features of the way in which epistemic vocabulary interacts with other epistemic vocabulary. If we start by studying the behavior of simple nested epistemic modals, we may naturally build a theory that explains the more complicated behavior of epistemic modals under disjunction and over indicative conditionals, and even the puzzling effects of embedding epistemic vocabulary in classically valid arguments. In Section 1, I make unifying observations about the suggestive behavior of epistemic vocabulary in each of these contexts, extracting several desiderata for semantic and pragmatic theories.

In Sections 2 and 3, I develop a semantics for epistemic vocabulary. This semantics constitutes a rather dramatic alternative to standard truth conditional theories, as it assigns sets of probability measures rather than sets of worlds as semantic values. I aim to demonstrate that what my theory lacks in conservatism, it makes up for in strength. In Section 4, I argue that combined with a novel pragmatics, my semantic theory can account for the distinctive linguistic behavior observed in Section 1. The theory I defend thereby addresses several challenges raised in recent literature. For instance, the theory answers concerns about epistemic modals under disjunction raised in Schroeder 2012. The theory also explains why epistemic vocabulary produces invalid instances of classically valid arguments, shedding light on important puzzles raised for constructive dilemma arguments in Kolodny & MacFarlane 2010 and modus tollens arguments in Yalcin 2012a.

1   Data for a theory of epistemic vocabulary

A careful examination of the behavior of epistemic modals yields several desiderata for a theory of epistemic vocabulary. A few of these desiderata have been discussed elsewhere, usually in the context of puzzles concerning epistemic modals. A number of the desiderata make trouble for extant semantic theories. The literature on epistemic modals is so vast that it would be impractical to argue against every alternative to my preferred theory here. For considerations of space, I set aside the possibility of resuscitating the standard truth conditional semantics for epistemic vocabulary, since persuasive arguments against that semantics have been discussed at length elsewhere.[1] I point out potential challenges for other prominent theories in

---

1 For instance, see the implications of triviality results discussed in Edgington 1995, the discussion of the subject matter of indicative conditionals in Bennett 2003, the "speaker

passing, but the main focus of this paper is the exposition and development of a positive case for my own theory.

1.1 Nested epistemic vocabulary

Nested epistemic vocabulary occurs in many forms in ordinary conversation. For example, suppose Alice and Bob are both candidates for certain job positions. We may naturally talk about Bob using epistemic adjectives under epistemic operators:

(5)   Alice is a likely hire, and Bob might be a likely hire.

(6)   Alice is a possible hire, and Bob is probably also a possible hire.

And we could further spell out the above observations as follows:

(7)   It is likely that we will hire Alice, and we might also be likely to hire Bob.

(8)   We might hire Alice, and it is probably the case that we might hire Bob too.[2]

Both epistemic modals and epistemic comparative adjectives can occur in the scope of indicative conditionals, and vice versa:

(9)   If they did not hire Alice, they are more likely to have hired Bob than Carl.[3]

(10)   It is more likely than not that the vase broke if he dropped it on concrete.

---

inclusion constraint" in Egan, Hawthorne & Weatherson 2005 and Weatherson 2008, the case of the missing car keys in Swanson 2006 and von Fintel & Gillies 2011, the eavesdropping cases in Egan 2007, the discussion of embedding behavior in Yalcin 2007, the discussion of inference patterns in Yalcin 2010, the discussion of assertability and disagreement in Yalcin 2011, and the discussion of retraction and disputes in MacFarlane 2011.

2 It cannot be taken for granted that both modals in these constructions are genuinely epistemic. However, in the next section of this paper, I give several arguments against the claim that one can always provide embedded modals with non-epistemic interpretations.

3 Hacquard & Wellwood 2012 give attested cases of epistemic vocabulary in indicative antecedents, while arguing that pragmatic considerations may limit the distribution of epistemic vocabulary in indicative antecedents and similar linguistic contexts.

In addition, there are well-known examples of right-nested and left-nested indicatives:

(11)   If a Republican wins the election, then if it's not Reagan who wins it will be Anderson. (McGee 1985, 462)

(12)   If the cup broke if it was dropped, it was fragile. (Gibbard 1981, 237)

And finally, there are attested uses of nested epistemic expressions occurring in short succession:

(13)   She could not but think [that] Wentworth was not in love with either. They were more in love with him; yet there it was not love. It was a little fever of admiration; but it might, *probably must*, end in love with some.[4]

(14)   The time is now near at hand which *must probably* determine, whether Americans are to be, Freemen, or Slaves.[5]

In wordy constructions such as (7) and (8) as well as condensed constructions such as (13) and (14), we are intuitively using nested epistemic modals to say something different from what we would use single modals to say. For example, intuitively (5) says something different about Bob than it says about Alice:

(5)   Alice is a likely hire, and Bob might be a likely hire.

To take another simple example, (15) intuitively says something different about Bob from either (16) or (17):

(15)   It is probably the case that Bob is a possible hire.

(16)   It is probably the case that Bob is a hire.

(17)   Bob is a possible hire.

In particular, our judgments suggest that (15) is weaker than either (16) or (17). Believing (16) is intuitively sufficient reason to bet at even odds that we will hire Bob, whereas merely believing (15) is not. Evidence for the semantic difference between (15) and (17) comes from direct intuitions about what we

---

4 Austen 1818, p. 55; italics added.
5 George Washington's address to the Continental Army before the Battle of Long Island, 27 August 1776; italics added.

use these sentences to talk about. In particular, nested epistemic modals are often used when you do not yet have some settled opinion on some question. If you say that Bob is a possible hire, it sounds as if you know that we might hire Bob. By contrast, if you merely say that it is *probably* the case that Bob is a possible hire, it sounds as if you have not yet settled on an opinion about Bob. Either Bob is a possible hire, or he isn't, and you are more inclined to side with the former opinion.

Relatedly, subjects sometimes report that they can easily make sense of nested epistemic modals by imagining that the speaker has several sources of information about their prejacent, and she is not sure which source she should trust. For instance, suppose we survey several equally informed experts about whether we might hire Bob. If most say that we might hire Bob and just a couple of experts disagree, then it is natural to form the opinion that it is probably the case that we might hire Bob. And analogous generalizations hold for other uses of nested epistemic modals. To comment on the example (14) above: if you say that some battle must *probably* be decisive, it sounds as if whatever settled opinion you may eventually have about the importance of the battle, you will settle on an opinion according to which the battle is probably decisive. It is easy to make sense of this state by imagining that you have several sources of information about whether the battle will be decisive, where each source agrees that the battle is at least more likely than not to be decisive.

According to naïve orthodoxy, when someone utters a declarative sentence, you should add its content to your stock of full beliefs. But as theorists have developed alternatives to full belief models of mental states, many have argued that what we say reflects what we think according to these more intricate models. For instance, some have claimed that epistemic modals are used to communicate partial beliefs.[6] At a first glance, it may appear that sentences containing nested epistemic modals are used to communicate even more intricate mental states. In particular, according to *imprecise credence* models, you are associated with multiple probability measures when you are unsettled as to how likely various propositions are, exactly as you might be when you are unsure what source of information you should trust. Rothschild 2012 argues that epistemic modals are used to communicate these sorts of imprecise credal states. The theory I develop does not model subjects as having imprecise credences. But whether or not we adopt the sort

---

6 See Section 2 for further discussion, and see Swanson 2012 for a recent catalog of relevant literature.

of semantics Rothschild defends, it is important that our theory account for intuitive judgments that can naturally be taken to support that proposal. In other words, the above discussion highlights an important goal for any theory of epistemic vocabulary. This is our first desideratum: our theory should explain why nested epistemic modals signal that different opinions about some subject are in play. Relatedly, our theory should explain why we sometimes easily make sense of embedded modals by imagining that a speaker bases her opinions on multiple sources of information.

A second desideratum for our theory of epistemic vocabulary is inspired by Yalcin 2007. Yalcin points out that our theory of epistemic possibility modals should explain why conjunctions of pairs of sentences such as (18) and (19) sound bad, and why such conjunctions continue to sound bad when embedded under indicative supposition, as in (20) and (21):

Some detectives are discussing the identity of a certain masked murderer.

(18)   It is not John.

(19)   It might be John.

(20)   #Suppose it is not John and it might be John.

(21)   #If it is not John and it might be John…

Along the same lines, note that not only is it bad to assert (18) and (19) together, but it is difficult to imagine a single circumstance in which you could be equally correct in uttering either of these sentences individually. If you would be correct in uttering (18) in some circumstance, then it is difficult to imagine how you could simultaneously be just as correct in uttering (19).

In this last respect, (18) and (19) stand in striking contrast to a similar pair of sentences, namely sentences that resemble (18) and (19), but where the embedded sentence is replaced with a sentence containing epistemic vocabulary:

(22)   It is not the case that it is probably John.

(23)   It might be the case that it is probably John.

It is possible to imagine a single circumstance in which you could correctly utter either (22) or (23). For instance, suppose you simply cannot make up your mind about how likely it is that the masked murderer is John. A few experts believe it is probably John, but a majority of experts believe it is probably Mary. In this case, you might correctly use (22), insofar as you

would side with the majority of experts if forced to choose one suspect. But you might also correctly use (23), insofar as you refuse to simply ignore the minority expert opinion. Here different frames of mind are relevant to your imagined utterances: (22) reflects your opinion after collating the advice of your expert advisors, while (23) reflects the fact that you are still not sure which experts you should trust. And of course, neither frame of mind vindicates the assertion of both sentences:

(24)   #It is not the case that it is probably John and it might be the case that it is probably John.[7]

These judgments yield a second desideratum for our theory of epistemic vocabulary: our theory should explain why in certain circumstances, we could correctly utter either (22) or (23), though we could not correctly utter their conjunction.

A third desideratum comes from a final observation about nested modals, namely that the strength of the outer modal often reflects the weight of your evidence and resilience of your opinion about the prejacent of the inner modal. For example: suppose that Liem likes wearing green shirts. His dad Eric has observed the color of his shirt on 800 consecutive days. Liem was wearing green on 500 of those days. His friend Madeleine has observed the color of his shirt on 8 consecutive days. Liem was wearing green on 5 of those days. Suppose that Eric and Madeleine have not yet seen what Liem is wearing today. Both Eric and Madeleine have .625 credence that Liem is wearing green, and both might guess that Liem is probably wearing green. But it seems more appropriate for Madeleine to assert (25) or (26), whereas Eric is intuitively licensed in asserting (27):

(25)   It might be probable that Liem is wearing green.

(26)   In fact, I'm fairly confident that he is probably wearing green.

(27)   Liem is definitely likely to be wearing green.

The assertability of (27) tracks two differences between Eric and Madeleine. Eric bases his credences about Liem on more evidence. In addition, his high credence that Liem is wearing green is more resilient. Joyce 2005 argues that in a number of evidential situations, "weight of evidence manifests itself

---

7 A less stilted but equally infelicitous version of the sentence: 'John isn't a probable killer and might be a probable killer'.

in the resilience of credences in the face of new data" (p. 166). In the above situation, both evidential weight and credal resilience are manifested in the strength of the modal that embeds (28):

(28)   Liem is probably wearing green.

Suppose you have a relatively uninformed hunch that Liem is probably wearing green. In other words, suppose that your high credence that Liem is wearing green is not justified by much evidence. Then you are intuitively licensed in asserting (25), but not (27). As you acquire more and more evidence, your high credence that Liem is wearing green will become more and more resilient, and you may embed (28) under stronger and stronger epistemic modals. Hence our third desideratum: our theory of epistemic vocabulary should explain this intuitive connection between nested modals, evidential weight, and credal resilience.

All three of the above desiderata pose challenges for several extant theories of epistemic modals. For example, consider the following standard dynamic semantic entries for epistemic possibility and necessity modals:[8]

$$c[\Diamond \phi] = \{w \in c \colon c[\phi] \neq \emptyset\}$$
$$c[\Box \phi] = c \setminus \{w \in c \colon (c \setminus c[\phi]) \neq \emptyset\}$$

From these definitions, we can derive the characteristic axioms of S5. Hence, according to this semantics, any string of possibility and necessity modals is equivalent to its innermost modal. Some dynamic semanticists explicitly embrace this result, claiming that "embedding an epistemic modal under another epistemic modal does not in general have any interesting semantic effects" (Willer 2013, 12). The same result holds for a prominent competitor of the dynamic semantic proposal, namely the semantics defended in Yalcin 2007. As Yalcin explains:

> Iterating epistemic possibility operators adds no value on this semantics... This may explain why iterating epistemic possibility modals generally does not sound right, and why, when it does, the truth-conditions of the result typically seem equivalent to $\Diamond \phi$. I will generally ignore iterated epistemic modalities.
>
> (p. 994)

---

8 For canonical instances of semantic proposals along these lines, see Stalnaker 1970, Veltman 1996, Beaver 2001, von Fintel & Gillies 2008a, and Willer 2013.

It is difficult to see how semantic proposals in this spirit could successfully explain the pervasive nature of nested modals, much less account for their distinctive behavior.

1.2 Against contextualist re-interpretations of nested epistemic vocabulary

The most substantive recent attempt at a more responsive semantics for nested epistemic modals appears in Yalcin 2009, where Yalcin admits that sometimes nested modals do

> ...allow for coherent interpretations not equivalent to corresponding expression with the most narrow modal. The latter case is not provided for by the above semantics. In such cases I would be inclined to appeal to tacit shifting of the information state parameter, akin to free indirect discourse.   (p. 21)

For further elaboration, we are directed to the following passage in Yalcin 2007:

> Interpretation may involve a tacit shift in the information parameter...to the target state of information for the context. Aside from Gricean considerations of charitable interpretation, it is not obvious whether general principles are involved in the interpretation of such tacit shifts.   (p. 1013)

It is difficult to know exactly what is intended by these brief suggestions, and hence my arguments so far may be understood as an invitation to develop these suggestions into a theory that satisfies the desiderata given above.

A natural development of these suggestions might be to say that in any sentence where nested modals occur, the prejacent of the outer modal receives the same boring sort of semantic value as any simple declarative sentence. For instance, one might assimilate sentences such as (27) with sentences about particular probability functions, such as (29) or (30):

(27)   It is almost certainly the case that Liem is probably wearing green.

(29)   It is almost certainly the case that the objective chance that Liem is wearing green is high.

(30)   It is almost certainly the case that my epistemic probability that Liem is wearing green is high.

However, there are many reasons to be skeptical of this approach. Recall that recent literature has provided a host of reasons to reject the claim that the prejacent (28) is equivalent to some simple declarative sentence like (31) or (32):

(28)   Liem is probably wearing green.

(31)   The objective chance that Liem is wearing green is high.

(32)   My epistemic probability that Liem is wearing green is high.

The crucial dialectical point to appreciate is that analogous concerns tell against the equivalence of these same sentences when they are embedded under epistemic vocabulary. For example, it is suspiciously difficult to say exactly what salient probability function (27) is talking about. In the case described above, Eric can utter (27). But he cannot utter (29), because Eric knows that the objective chance that Liem is wearing green is either 0 or 1, and Eric is not almost certain of the latter. Madeleine cannot utter (27). But she can utter (30), because she knows that her inductive evidence confirms the claim that Liem is wearing green. Hence neither (29) nor (30) accurately paraphrases (27).

  Furthermore, eavesdroppers may explicitly target the prejacent of (27) and correctly evaluate it relative to their epistemic situation. For instance, if I have just seen Liem wearing a red shirt and I overhear Eric utter (27), it would be pedantic but nevertheless acceptable for me to say:

(33)   That isn't almost certain; it's just *false*. It's not the case that Liem is probably wearing green — he is wearing red.

A notorious dilemma for truth-conditional accounts replays itself here. If Eric was using 'probably' just to talk about his own evidential situation, then I am not licensed in saying 'it's false' in judging the prejacent of (27). On the other hand, if Eric was using 'probably' to talk about some evidence that included my evidence, then he was not licensed in uttering (27) to begin with.[9]

---

9 This is just the first step in an involved dialectic. For further discussion of eavesdropping arguments against truth-conditional accounts of epistemic vocabulary, see Egan, Hawthorne & Weatherson 2005, Egan 2007, Hawthorne 2007, von Fintel & Gillies 2008b, Knobe & Yalcin 2015, and MacFarlane 2011.

In fact, nearly every argument against a uniform truth conditional theory of all epistemic modals yields an analogous argument against a uniform truth conditional theory of all embedded epistemic modals. Bennett 2003 may argue that any alleged paraphrases of (27) fail to capture its intuitive subject matter, for instance. Bennett argues that when someone utters an indicative conditional,

> …common sense and the Ramsey test both clamour that [she] is not assuring me that her value for a certain conditional probability is high, but is assuring me of that high value…She aims to convince me of that probability, not the proposition *that* it is her probability. (p. 90)

Yalcin 2011 adds that the *reasons* that I give in support of my utterance 'it might be raining' concern the first-order proposition that it is raining, rather than any contextually determined body of evidence. Both Bennett and Yalcin could complain that (27) intuitively concerns Liem, rather than any contextually determined body of evidence. Another challenge comes from Yalcin 2007. If embedded modals are always interpreted relative to some salient probability function, then we lack an explanation for the infelicity of sentences such as:

(34)   #Probably, it is raining and might not be raining.

(35)   #It is unlikely that it is both raining and might not be raining.

(36)   #It might be that it is both raining and might not be raining.

These judgments are not accommodated by expressivist, relativist, or dynamic theories that resort to assigning simple semantic contents to embedded modal constructions.

In addition, it is worth noting that if we reinterpret the prejacent of (27) as having straightforward truth conditions, we are still left with the problem of interpreting (37b) in the following dialogue:

(37)   a.   *David:* Is Liem probably wearing green?
       b.   *Eric:* Almost certainly.

Familiar arguments challenge the claim that the unembedded (37a) has straightforward truth conditions. Furthermore, it is difficult to see why Gricean considerations should demand that we interpret (37a) as containing

free indirect discourse or a tacitly shifted information parameter. Hence it seems we must find some way of interpreting (37b) without appealing to such strategies. One would expect the resulting understanding of (37b) to provide some similar understanding of (27), namely an alternative semantics that recognizes that 'Liem is probably wearing green' need not express a possible worlds content, whether it is embedded in a question or under further epistemic vocabulary. To sum up: it is not obvious that extant semantic theories can explain the behavior of nested epistemic modals. A natural way of developing potential explanations on behalf of recent expressivist, relativist, and dynamic theories meets with several challenges. Hence the behavior of nested epistemic modals should motivate us to look for alternative semantic theories.

1.3 Epistemic vocabulary under disjunction

A fourth desideratum for our theory of epistemic vocabulary is inspired by Schroeder 2012. Schroeder argues that a semantic theory should *not* predict that you can assert a disjunction only if you can assert one of its disjuncts, even in special cases where disjuncts are stipulated to be governed by wide-scope epistemic modals. Schroeder points out several reasons why this prediction would be bad. Here is one example:

> Last night Shieva calls me to express frustration with the paper that she is working on, and tells me that if she hasn't finished by this morning, she's going to consult her magic 8-ball about whether to give up and follow its advice. Since I know that most of the answers on her magic 8-ball are positive, when I recall our conversation from last night, I conclude that either Shieva finished her paper by this morning, or she probably gave up.
>
> (pp. 21–2)

In this case, the speaker can correctly assert 'Shieva finished or probably gave up' without being able to assert either disjunct. Similarly, you can correctly assert (38) about the result of throwing a fair die, without being able to assert either disjunct:

(38)   It is less than four or probably even.

In this respect, disjunctions embedding epistemic vocabulary are just like ordinary disjunctions of simple sentences. In fact, asserting a disjunction usually implicates that you are not in a position to assert either disjunct. There is something especially peculiar about disjunctions embedding epistemic vocabulary, though. Even if you can deny one disjunct and you cannot assert the other, you may still be able to assert the entire disjunction. For instance: you can assert (38) even though you can deny the second disjunct by itself, and you cannot assert the first. This does not hold for disjunctions without epistemic vocabulary. If you can deny one half of a simple disjunction, then disjunctive syllogism ordinarily proves that the remaining disjunct is equivalent to the entire disjunction, so one is not assertable without the other. This brings us to our fourth desideratum: our theory should explain this surprising difference between simple disjunctions and disjunctions containing epistemic vocabulary.

A semantics for 'or' is missing from Yalcin 2007, 2011, 2012a, and related papers. Hence the relevant challenge for Yalcin is to state a semantics that predicts the behavior just described.[10] Substantially more progress has been made on disjunction in the dynamic semantics literature. In fact, a number of dynamic accounts of disjunction satisfy our fourth desideratum. According to these accounts, natural language disjunction is not commutative. Roughly speaking, the second half of a disjunction is not interpreted relative to a global context, but rather relative to a local context that has been updated with the negation of the first disjunct. This sort of account aims to give a uniform explanation of the local interpretation of 'probably' in (38) and local satisfaction of licensing conditions for pronouns in disjunctions such as the following famous example from Roberts 1989:

(39)   Either there is no bathroom in this house, or it is in a funny place.

Just as the licensing conditions for 'it' in (39) are satisfied in a local context where the first disjunct is false, values of contextual parameters in the second disjunct of (38) are provided by a local context where the first disjunct is false. This explains why you may assert (38) even when you can deny its second disjunct uttered in isolation. The disjunction is felicitous because its second disjunct is acceptable in all contexts where the negation of the first disjunct is given.

---

10 Schroeder extrapolates a semantics for 'or' from Yalcin 2007 and criticizes that semantics for validating 'or' exportation.

This dynamic account predicts that natural language disjunction is not commutative, and fans of this account often claim this predicted failure of commutativity as a benefit. For instance, they claim that a semantics for natural language disjunction should entail that (40) sounds bad even though (38) sounds fine:

(38)   It is less than four or probably even.

(40)   It is probably even or less than four.

However, it is not clear that we should want our semantics to predict this difference between (38) and (40).[11] For instance, there are a number of contexts in which (40) seems just as good as (38), namely contexts in which certain partitions of logical space are salient. Consider the following case:

> Alice just rolled a fair die and hid it under a cup in front of me. I see a blue cup and a red cup. The die is under the blue cup if it landed on a four, five, or six. The die is under the red cup if it landed on a number less than four.
>
> Bob offers me a pair of bets. For one dollar, he will sell me a bet that pays five dollars if the die landed on an even number. For another dollar, he will sell me a bet that pays five dollars if the die landed on a number less than four. I am very risk averse, and I do not always bet to maximize expected returns. But staring first at the blue cup and then at the red cup, I judge that I would be comfortable accepting both bets, since, as I put it, "either it is probably even, or less than four."

The circumstances of the above case call attention to a certain partition of logical space: either the die landed on a number less than four, or it landed on a higher number. Against this background, my utterance of (40) seems perfectly correct.[12]

---

11 The commutativity of disjunction is controversial even among advocates of dynamic semantic theories. For instance, Schlenker 2009 and Rothschild 2011 both provide theories according to which disjunction is commutative; their accounts are sympathetic with my discussion of the fifth desideratum.

12 Some readers may find it difficult to evaluate the artificial speech described above, especially since the salience of an objective chance function may introduce noise in our judgments. The essential point of the present discussion is that contextual cues may make certain readings

In fact, some disjunctions like (40) sound fine without heavy contextual cues. For instance, you can assert any of the following disjunctions, even if you can deny the first disjunct and cannot assert the second:

(41)   It's either unlikely he was being honest with you, or he just wanted you to think that he was lying.

(42)   The next United States president will either almost certainly attempt to repeal a lot of Barack Obama's policies, or they will be a Democrat with more liberal views than Obama has.

(43)   John is probably playing baseball, or it has been raining all afternoon.

These disjunctions seem to mean the same thing regardless of the order in which their disjuncts are uttered. In fact, they might just as well be written with their disjuncts arranged in a circle, without detriment to our ability to understand or evaluate them. This yields a fifth desideratum for our theory of epistemic vocabulary: our theory should explain why disjunctions such as (40) sound infelicitous in some contexts and felicitous in others. And our theory should explain why reversing disjunct order does not affect the interpretation of disjunctions in contexts where they sound felicitous.

This fifth desideratum should give us pause before we endorse a semantic theory that explicitly entails that natural language disjunction is not commutative. Furthermore, the above dynamic explanation for why we can assert (38) seems insufficiently general, since it does not explain why we can sometimes assert (40)–(43). The dynamic proposal outlined above says that we can sometimes assert a disjunction like (38) when its second disjunct is deniable and its first disjunct is unassertable. But (40)–(43) are all sometimes assertable even when their *first* disjuncts are deniable and their *second* disjuncts are unassertable. According to the dynamic explanation, (38) is felicitous because its second disjunct is acceptable in all contexts where the negation of the first disjunct is given. But for any of (40)–(43), the second disjunct is not acceptable even in contexts where the negation of the first disjunct is given. For example, the negation of the first disjunct of (40) is already given in an ordinary context where a fair die is rolled, but the second disjunct of (40) is not acceptable in that context:

(40)   It is probably even or less than four.

---

of epistemic vocabulary available. See Section 4.5 for more natural illustrations and a more detailed defense of this point.

To sum up: several observations raise challenges for several extant dynamic semantic accounts of the assertability of disjunctions. In particular, differences in the assertability of (38) and (40) seem sensitive to contextual factors, such as the salience of various alternative sets. This should motivate us to doubt theories that derive differences in assertability from context-insensitive semantic rules. Pragmatic theories are better designed to account for the distinctive behavior of disjunctions embedding epistemic vocabulary.

1.4 Epistemic vocabulary over indicatives

A sixth desideratum for a theory of epistemic vocabulary is inspired by an example in chapter 9 of Lycan 2001, which itself builds on a related discussion of subjunctive conditionals in Slote 1978. Consider the following case:

> Jill is standing on the roof of your office building. The local fire department occasionally hangs a net along the roof to protect workers doing construction. The net is strong enough to safely catch anyone who falls off the building. Just a few hours ago, you happened to notice that there was no net along the roof. As a result, you do not believe that Jill is going to jump off the roof. Jill is a thrill-seeker who might jump into a net for fun, but she definitely does not have a death wish. And without a net, anyone who jumped off the roof would surely fall to the ground and die.

On the one hand, since you believe that there is no net along the roof, you are intuitively justified in asserting:

(44)   Probably, if Jill jumps off the building, she will die.

On the other hand, you are confident that Jill does not have a death wish. If you were informed that Jill jumped off the building, you would immediately conclude that the local fire department must have installed a net since you last checked the roof. With that information in the front of your mind, you are intuitively justified in denying (44) and asserting:

(45)   Probably, if Jill jumps off the building, she will live.

To make these observations more vivid, suppose someone asks you whether there is a net along the roof of the building. They may well know that you

promised the fire department that you wouldn't go around telling people whether or not there was a net along the roof, but they may still persist in pestering you for information. It is intuitively fine for you to respond:

(46)   I cannot answer your questions directly. But I can tell you this much: it is really likely that if Jill jumps off this building, she will die.

On the other hand, suppose someone asks you whether you believe that Jill is suicidal. Again, they may well know that you promised Jill that you wouldn't go around telling people about her mental state, but they may persist in pestering you for information. Suppose that it is common ground that anyone suicidal would simply cut away any safety net and jump off the building in question. It is intuitively fine for you to respond:

(47)   I cannot answer your questions directly. But I can tell you this much: it is really likely that if Jill jumps off this building, she will live.

Hence the assertability of (44) does not depend only on your opinions about Jill and the net, which we may stipulate are the same when you utter (46) and (47). It must also be sensitive to some factor that varies between these contexts of utterance.

As with many other examples we have considered, you are considering different questions in these different contexts, and which question you are considering seems relevant to which utterances are felicitous. Suppose you are considering the question of whether there is a net along the roof. Then since you believe that there is probably no net, you may say that it is probably the case that if Jill jumps from the roof, she will die. Suppose you are considering the question of whether Jill is suicidal. Then since you believe that she is probably not suicidal, you may say that it is probably the case that if Jill jumps from the roof, she will live. The sixth desideratum: our theory of epistemic vocabulary should explain this variation in the assertability conditions of (44).

In many extant theories of epistemic vocabulary, there is no obvious mechanism for explaining this variation. The semantic values for 'probably' and 'if' given in Veltman 1996 and Yalcin 2012a do not depend on contextually determined parameters. An advocate of these semantic proposals might attribute variation in the interpretation of (44) to scope ambiguity. At the level of logical form, 'probably' might take scope over the entire indicative conditional in (44) or just over its consequent. But this does not seem like a

plausible explanation of the behavior of (44), since context not only affects our interpretation of (44), but also our interpretation of the unembedded indicative conditional (48):

(44)   Probably, if Jill jumps off the building, she will die.

(48)   If Jill jumps off the building, she will die.

The unembedded conditional is borderline assertable when we are focusing on whether there is a net along the roof, but definitely unassertable when we are focusing on whether Jill is suicidal. These judgments suggest that the interpretation of the indicative itself depends on contextually determined parameters.

A related challenge arises when we embed sentences like (44) in indicative conditionals. If we are talking about whether there is a net, you can correctly assert:

(49)   If it is probably the case that Jill will live if she jumps, then there is a net.

If we are talking about whether Jill is suicidal, you can correctly assert:

(50)   It is probably the case that Jill will live if she jumps.

However, you can never correctly assert:

(51)   There is a net.

These judgments make trouble for certain semantic theories. Several dynamic and expressivist theories say something roughly like the following: you believe a sentence when your credal state accepts it. And an information state accepts a conditional when the closest state that accepts its antecedent also accepts its consequent. Since you believe (50), your actual credal state accepts the antecedent of (49). Hence your actual credal state is the credal state closest to yours that accepts that antecedent. Since you believe the conditional (49), we should conclude that your actual credal state also accepts its consequent (51). But this conclusion seems clearly false.[13]

---

13 In order to keep my discussion as general as possible, I will not use this formula to construct objections for particular theories. The interested reader should combine the discussion of attitude verbs in Section 7 of Yalcin 2007 with the semantics for 'if' and 'probably' in the appendix of Yalcin 2012a. For dynamic theories, combine the standard dynamic semantics for attitude verbs in Heim 1992 with the dynamic semantics for 'if' and 'probably' developed

The complex conditional (49) gives rise to our seventh desideratum: our theory should explain its assertability conditions. This is not a trivial endeavor. First, our theory must assign semantic contents to indicatives whose antecedents embed both graded epistemic vocabulary and other indicatives. Second, our theory must explain how your beliefs can support asserting (49) in some contexts and (50) in others, without ever supporting (51). These facts intuitively depend on the context sensitivity of (49) and (50), and relevant contextual factors intuitively include facts about what questions are salient when each is uttered.

1.5 Epistemic vocabulary in classically valid arguments

The seventh desideratum also directs us toward one final category of useful observations. If you believe both (49) and (50), it might seem that you could apply *modus ponens* and infer that there is a net along the roof. But you are not licensed in believing that there is a net along the roof. The final three desiderata concern instances of classically valid argument forms that seem invalid in virtue of containing epistemic vocabulary.

Suppose Carlos has rolled a fair die without telling us how it landed. A fair die has three low numbers and three high numbers. Suppose we are considering the following argument about the number Carlos rolled:

(52) a. If it is low, it is probably odd.
      b. It is not probably odd.
      c. Hence: it is not low.

This argument seems like an instance of *modus tollens*. But it also seems invalid. The first premise seems correct, since 2 out of 3 of the low numbers are odd. The second premise seems correct, since it is just as likely that an even number was rolled as an odd number. But these premises do not justify our accepting the conclusion, since we have no idea whether a low number was rolled. Several authors have made similar observations about apparent instances of *modus tollens* containing epistemic modals.[14] This raises a puzzle: should we say that (52) is not an instance of *modus tollens*,

---

in Section 4 of Gillies 2004, Section 10 of Gillies 2010, or the appendix of Yalcin 2012a, replacing "closest credal state to yours that accepts the antecedent" with "result of updating your credal state on the antecedent" in my discussion above.

14 For related discussion, see Carroll 1894, Veltman 1985, Cantwell 2008, and especially Yalcin 2012a.

that (52) is valid, or that some instances of *modus tollens* are not valid? This brings us to our eighth desideratum: our theory of epistemic vocabulary should solve this puzzle. At a minimum, our theory should come equipped with a notion of consequence that yields a verdict about whether (52) is valid. And whether or not it is valid, our theory should predict the apparent invalidity of instances of *modus tollens* containing epistemic vocabulary.

Here is another apparently invalid argument about the number rolled:

(53)   a.   If it is low, it is probably odd.
       b.   If it is high, it is probably even.
       c.   It is either low or high.
       d.   Hence: either it is probably odd or probably even.

Kolodny & MacFarlane 2010 discuss similar arguments, including the following:

(54)   a.   Either the butler did it or the nephew did it.
       b.   If the butler did it, the murder must have occurred in the morning.
       c.   If the nephew did it, the murder must have occurred in the evening.
       d.   Hence: either the murder must have occurred in the morning or it must have occurred in the evening.

These arguments seem like instances of constructive dilemma. But they also seem invalid. For instance, just as it seems incorrect to say that the number rolled is probably even, it seems incorrect to say it is probably odd. So in the absence of any special contextual cues, it seems incorrect to say that the number rolled is either probably even or probably odd. It is neither probably even nor probably odd, but just as likely to be one or the other. This brings us to our ninth desideratum: our theory should say whether (53) is valid. And whether or not it is valid, our theory should predict the apparent invalidity of instances of constructive dilemma containing epistemic vocabulary.

Similar problems arise not just for *modus tollens* and constructive dilemma, but also for disjunctive syllogism:

(55)   a.   It is low or probably even.
       b.   It is not probably even.
       c.   Hence: it is low.

And contraposition of indicative conditionals:

(56)   a.   If it is low, it is probably even.

            b.   Hence: if it is not probably even, it is not low.

Furthermore, it seems entirely appropriate to give similar explanations for the apparent invalidity of these inferences. Kolodny & MacFarlane 2010 and Yalcin 2012a, for instance, defend semantic theories according to which each of the relevant inference rules is literally invalid. In fact, Kolodny and MacFarlane go so far as to say that *modus ponens* itself is an invalid rule of inference.

    Anyone rejecting classically valid inference rules bears the burden of explaining why we successfully use them in ordinary reasoning. The easiest way to discharge this burden is by proving that the rules are indeed valid when restricted to premises of a certain form. At a minimum, setting aside complications involving adverbs of quantification, it seems our theory should predict that arguments are valid when they contain no epistemic vocabulary at all. This condition raises an important question, namely exactly which arguments containing epistemic vocabulary are valid.

    Kolodny & MacFarlane 2010 defend inferences involving conditionals whose consequents do not contain any epistemic vocabulary. However, some inferences involving conditionals whose consequents contain epistemic vocabulary are intuitively valid as well. For instance, Yalcin 2012a suggests that the following inference is valid:

(57)   a.   If the marble is big, then it might be red.

            b.   It is not the case that it might be red.

            c.   Hence: it is not big.

In addition, some probabilistic inference rules are intuitively valid, and some of those rules govern indicatives with consequents embedding epistemic vocabulary. In fact, we just considered inferences of this sort in Section 1.4. The following inference licenses my saying (58c) when discussing whether there is a net along the roof:

(58)   a.   Probably, there is no net along the roof.

            b.   If there is no net along the roof, then if Jill jumps, she will die.

            c.   Hence: probably, if Jill jumps, she will die.

And the following licenses my saying (59c) when discussing whether Jill is suicidal:

(59)   a.   Probably, Jill is not suicidal.
       b.   If Jill is not suicidal, then if Jill jumps, she will live.
       c.   Hence: probably, if Jill jumps, she will live.

This brings us to our tenth and final desideratum for a theory of epistemic vocabulary. Insofar as our theory says that standard inference rules are generally invalid, it should explain why substantial classes of restricted rules appear to be genuinely valid. In particular, our theory should explain why (57), (58), and (59) are apparently valid, even though these inferences are riddled with epistemic vocabulary.

2   A basic semantics for epistemic vocabulary

Before stating specific semantic entries, it will be helpful to outline the basic idea of the semantic theory itself. Recall that in a certain context, you may correctly describe the outcome of rolling a fair die by saying:

(40)   It is probably even or less than four.

The imagined context of (40) is somewhat contrived. In particular, the context is contrived to make a certain partition of logical space especially salient. The partition has two elements: either the number rolled is low, or it is high. As a result, there are also two kinds of salient credence distributions when you utter (40). First, you have conditional credences, conditional on the partition propositions. For example, you have higher than .5 credence that the number rolled is even, conditional on it being high. Second, you have a credence distribution over the partition propositions themselves. For example, you have .5 credence that number rolled is high. In other words, there are various opinions you might have after learning some information from the contextually salient partition. And on top of that, you have some opinions about the likelihood of each bit of information that you could learn.
   Here is a first pass at my semantics: the latter opinions are associated with higher modals, while the former are associated with embedded modals. For example, it would sound fine for you to say (60) in the context mentioned above:

(60)   It might well be that the number is probably even.

According to my semantics, that is roughly because you could learn *some* salient information — namely that the number rolled is high — confirming an

opinion that gives *most* of its credence to the number rolled being even. To take another example, suppose that you are torn between various ways of evaluating candidates for an academic position. It is not clear how to weigh teaching experience against research quality, for instance, and you are open to information that would decide this question in different ways. In spite of your indecision, you might say:

(61)   It must be the case that Bob is a possible hire.

According to my semantics, that is roughly because as far as your credences are concerned, *any* salient information would support an opinion that gave at least *some* credence to Bob being hired. Again, the embedded modal ('possible') is associated with your credences conditional on various propositions (about ways of evaluating candidates), while the higher modal ('must') is associated with your credences in those propositions themselves.

According to a traditional account of assertion, an assertion is "something like a proposal" (*cf.* Stalnaker 2010, p. 152), namely the proposal that the content of the assertion be added to the propositions taken for granted in the conversation. In a paradigmatic case of assertion, you believe a proposition, you assert some sentence with that proposition as its content, and as a result, I come to believe that same proposition. This model of assertion fits well with a certain model of our mental life, according to which full beliefs are the opinions we have and the opinions we want to share with each other. Meanwhile, theorists have developed alternate models of our mental life in which degreed beliefs play a central role. It is natural to wonder whether we can update our account of assertion to fit these more sophisticated models.

The updated account: an assertion is like a proposal, not about a proposition that you should believe, but rather about a property that your credences should have. It is still true that in a paradigmatic case of assertion, you have an opinion, you assert some sentence with that opinion as its content, and as a result, I come to have the same opinion. But the relevant opinions are degreed. In other words, having an opinion amounts to having credences with a certain property. The content of a declarative sentence is a property that credences can have. Formally, contents are sets of probability measures. In a paradigmatic case of assertion, when you assert a sentence with a certain content, I come to have a credence distribution that is contained in that content. For instance, you may assert a sentence whose content is the set of all measures that assign probability greater than .5 to the proposition that it is raining. On hearing your assertion, I will come to have more than

.5 credence that it is raining. Following Swanson 2006, we may conceive of the content of a sentence as a *constraint* on credences, namely the constraint that my credences generally end up satisfying on hearing your assertion of that sentence.

Sentences containing epistemic vocabulary are context sensitive. In other words, which set of measures is the content of a sentence depends on what context you are using the sentence in. In particular, context contributes partitions of logical space to the semantic values of such sentences. The contextually determined partitions make the contents of sentences more interesting. A second pass at the heart of my semantics: some asserted contents are straightforward constraints on credences, such as assigning greater than .5 credence to some particular proposition. But asserted contents can also constrain your credences to have more fine-grained properties. In particular, they can constrain the structure of your credences with respect to propositions in non-trivial contextually determined partitions. The content of a sentence containing nested epistemic modals will be a constraint having to do with your credences in those propositions, and also with your credences conditional on those propositions. Higher modals correspond to the former sort of constraint, while embedded modals correspond to the latter.

2.1   A semantics for logical operators

In addition to formal semantic entries, it will be useful to have some shorthand for saying what expressions mean. Let us say that your credences satisfy the constraint that a certain proposition *accepts that it is probably raining* just in case it is probably raining according to your credences conditional on that proposition, or in other words, just in case your conditional credences are contained in the content that it is probably raining. If context determines a partition of logical space, we can quantify over the members of that partition as if they were each identified with different people. For instance, given a contextually determined partition, let us say that your credences satisfy the shorthand constraint that *someone accepts that it is probably raining* just in case some proposition in the partition accepts that it is probably raining. In general, let us say your credences satisfy the constraint that *someone* accepts a particular content just in case there is *some* proposition in the partition such that your credences given that proposition are contained in that particular content. Your credences satisfy the constraint that *everyone* accepts a content just in case *every* proposition in the partition is such that

your credences given that proposition are contained in that content. And so on. Rather than always explicitly describing your credences conditional on propositions in a contextually determined partition, we have a handy shorthand that captures the sense in which your credences conditional on different partition elements often correspond to different states of opinion that you have not yourself decided between. In a rough sense, one may imagine the shorthand expressions 'someone' and 'everyone' as quantifying over different sides of yourself.[15]

Now for the semantics. In contrast with a number of extant theories, it is straightforward to start with a semantics for all basic logical operators, including natural language disjunction. For instance: your credences are contained in the content of a disjunction just in case every proposition in the corresponding contextually determined partition is such that your credences conditional on that proposition are contained in the content of one of the disjuncts. The semantic entries for 'and' and 'not' are predictable variants. In shorthand:

'S or T' means that everyone accepts that S or accepts that T.

'S and T' means that everyone accepts that S and accepts that T.

'not S' means that no one accepts that S.[16]

In more formal vocabulary:

$$\llbracket \mathbf{or}_i \rrbracket^c = [\lambda S.\lambda T.\{m \colon \forall p \in g_c(i), m|_p \in S \text{ or } m|_p \in T\}]$$

$$\llbracket \mathbf{and}_i \rrbracket^c = [\lambda S.\lambda T.\{m \colon \forall p \in g_c(i), m|_p \in S \text{ and } m|_p \in T\}]$$

$$\llbracket \mathbf{not}_i \rrbracket^c = [\lambda S. \quad \{m \colon \forall p \in g_c(i), m|_p \notin S\}]$$

A number of notes about the formal vocabulary are in order. The variable p ranges over sets of worlds, and m ranges over probability measures. The measure $m|_p$ is the result of conditionalizing the measure m on the proposition p. Let us stipulate that $\mathcal{S}$ is the semantic type of sets of measures.

---

15 In what follows, I often simplify my discussion by just talking about whether certain partition elements accept a certain constraint. It should be understood that strictly speaking, whether a proposition accepts a constraint is relative to a measure, meaning for instance that your credences may satisfy the constraint that someone accepts that it is probably raining, while my credences fail to satisfy this same constraint.

16 I use 'not' as shorthand for 'it is not the case that' and I treat this expression as an operator that occurs just before its argument, though ultimately one should allow many other expressions of sentential negation at surface structure. The analogous claims hold for 'might', 'must', and 'probably'.

In the above entries, the variables S and T range over values of type $\mathcal{S}$. The logical operators 'and' and 'or' have semantic values of type $\langle \mathcal{S}, \langle \mathcal{S}, \mathcal{S} \rangle \rangle$, whereas 'not' has a semantic value of type $\langle \mathcal{S}, \mathcal{S} \rangle$. For example, the content of a disjunction is a set of measures, as is the content of each disjunct.

Exactly which set of measures is the content of a disjunction depends on what partition context contributes to its content. Following Heim & Kratzer 1998, we say that every context c determines an assignment function g_c that specifies the values of all contextually determined variables. The value $g_c(i)$ is the contextually determined partition relevant to the semantic entry spelled out above. The shorthand expression 'everyone' corresponds to the formal expression '$\forall p \in g_c(i)$' which quantifies over propositions in that partition. In what follows, I use both shorthand and formal vocabulary, as the former allows me to make my arguments intuitive, while the latter allows me to make them precise.

In slightly less formal vocabulary, the semantic value of 'S or T' is the set of measures m satisfying the following condition: for any proposition p in the relevant contextually determined partition, $m|_p$ is either contained in the content of the first disjunct or in the content of the second disjunct. For example, recall that in some contexts where you have equal credence in each possible outcome of rolling a fair die, it sounds okay for you to say:

(40)   It is probably even or less than four.

As mentioned earlier, the sort of context that is hospitable for (40) makes a certain partition salient: either the number rolled is low, or it is high. According to your credences conditional on it being low, the number is less than four. According to your credences conditional on it being high, the number is probably even. Hence your credences satisfy the content of (40), namely that everyone in the contextually determined partition either accept that the number rolled is probably even or accept that it is less than four. In a nutshell: you believe (40), and that is why it sounds okay for you to say it.

This explanation is incomplete as it stands. For starters, a complete explanation requires identifying the content of each disjunct of (40) relative to the sort of context in question, so that we may prove that your conditional credences are contained in these contents. Appendix B.1 contains a complete explanation of why your credence is in the content of (40), and Section 2.4 contains further commentary. Another clarificatory note: the above semantic values are custom-made for logical operators embedding epistemic vocabulary. The theory I develop assigns more traditional semantic values to logical

operators elsewhere. The careful reader will observe that according to this theory, logical operators embedding epistemic vocabulary act essentially like epistemic vocabulary. This observation is implausible unless restricted to logical operators embedding epistemic vocabulary, so it is important to bear in mind that more traditional semantic values for logical operators will be revived in Section 3.

2.2   A semantics for epistemic possibility and necessity modals

Here are shorthand semantic entries for epistemic possibility and necessity modals:

'might S' means that someone accepts that S.

'must S' means that everyone accepts that S.

In more formal vocabulary:

$$[\![\textbf{might}_i]\!]^c = [\lambda S.\{m\colon \exists p \in g_c(i) \text{ such that } m|_p \in S\}]$$
$$[\![\textbf{must}_i]\!]^c = [\lambda S.\{m\colon \forall p \in g_c(i),\ m|_p \in S\}]$$

Having expanded our lexicon, we can outline a semantics for some nested epistemic modals. For example, (62) and (63) each mean that everyone accepts that someone accepts that we will hire Bob:

(62)   It is definitely the case that Bob might be the best candidate for the job.

(63)   It must be the case that Bob might be the best candidate for the job.

This shorthand calls attention to an important semantic feature: higher and lower epistemic modals need not be associated with the same domain of quantification. Both logical operators and modals have indices. Assignment functions map expressions with different indices to potentially different values. Hence unless expressions are co-indexed, context may contribute different partitions to their interpretation. For example, an utterance of (62) may contain modals that are not co-indexed:

(64)   It is definitely$_1$ the case that Bob might$_2$ be the best candidate for the job.

The semantic value of (64) is as follows

$$[\![(64)]\!]^c = \{m\colon \forall p \in g_c(1),$$
$$m|_p \in \{m'\colon \exists q \in g_c(2) \text{ such that } m'|_q \in [\![(65)]\!]^c\}\}$$

where (65) is the prejacent of the inner modal in (64):

(65)   Bob is the best candidate for the job.

For instance, in a context where (64) is uttered, it could be that the partition $g_c(1)$ contains propositions about what sorts of virtues matter when evaluating candidates, while the partition $g_c(2)$ contains propositions about which candidates have what sorts of virtues. In that sort of context, your credences would satisfy (64) just in case conditional on any proposition about what virtues matter, your credences satisfy the following condition: conditional on some proposition about which candidates have which virtues, Bob is the best candidate for the job.

For those especially attentive to syntactic representation: strictly speaking, our semantics could identify indexed variables as arguments of modals and logical operators, rather than indexing these expressions directly. For example, our formal semantic entry for 'must' could be as follows, where v ranges over partitions:

$$[\![\mathbf{must}]\!]^c = [\lambda v.\lambda S.\{m\colon \forall p \in v,\ m|_p \in S\}]$$

In that case, (62) would contain two covert pronouns:

(66)   It is definitely v_1 the case that Bob might v_2 be the best candidate for the job.

Here the pronouns v_1 and v_2 denote partitions relative to contexts, according to the familiar semantics for referential pronouns, namely $[\![v_i]\!]^c = g_c(i)$. The resulting semantic value of 'must v_i' matches the semantic value of 'must$_i$' given above. The reader may replace expressions of the latter sort with their kosher substitutes throughout.[17]

---

17 For simplicity, I will sometimes talk as if the contextually supplied partition is the value of a covert pronoun. But strictly speaking, I am neutral about the best syntactic implementation of my theory. Partee 1989 and Condoravdi & Gawron 1996 have given reasons to doubt that similar implicit arguments are best analyzed as the values of covert pronouns, and I will not evaluate their arguments in this paper.

2.3 A small detour: Advantages of constraining conditional credences

Recall from Section 1.1 that our use of nested epistemic modals fits naturally with the idea that sentences constrain imprecise credal states. This idea should seem even more compelling given all the shorthand just introduced. Suppose we model your mental state with a set of probability measures. In other words, suppose we model you as if you have an imaginary mental committee of subjects with precise credences. Then following Rothschild 2012, we could say that sentences constrain your mental committee members, rather than your conditional credences. If a sentence demands that everyone accepts a content, for instance, that could just amount to demanding that each committee member accept that content. In other words, my shorthand semantic entries for 'might' and 'must' seem like apt translations of the following alternative formal semantic entries:

$$[\![\mathbf{might}]\!] = [\lambda S.\{I \colon \exists m \in I \text{ such that } m \in S\}]$$
$$[\![\mathbf{must}]\!] = [\lambda S.\{I \colon \forall m \in I, \ m \in S\}]$$

Here the variable m ranges over precise credal states (*i.e.* probability measures) while S and I range over imprecise credal states (*i.e.* sets of probability measures). This proposal may appear to satisfy many desiderata given in Section 1. It is worthwhile to reflect on how my semantics differs from this proposal, and especially to notice that the imprecise credence proposal is deficient in two respects.

First, on the imprecise credence semantics stated above, embedding a sentence under 'might' or 'must' raises its semantic type. Each modal accepts sets of measures as inputs and delivers sets of imprecise credal states as outputs. That means that a sentence with a wide-scope 'might' or 'must' has the wrong semantic type to be embedded under another epistemic modal — a bad result, given our pervasive use of embedded modals. The most natural repair strategy requires that we model subjects as having not just imprecise credences, but more complicated mental states. In fact, very complicated mental states are required, since subjects commonly embed epistemic vocabulary under *embedded* epistemic vocabulary. For instance, recall that we have no trouble understanding (49):

(49)   If it is probably the case that Jill will live if she jumps, then there is a net.

And deeper embeddings seem perfectly intelligible, as long as the context is rich enough to supply the interpretations of relevant expressions. For instance, (49) sounds fine when you are trying to figure out whether there is a net along the roof of your office building. Suppose that the local fire department occasionally puts a trampoline instead of a net along the roof. Then we are not really licensed in saying that there is a net along the roof, given just that it is probably the case that Jill will live if she jumps. Instead, we should say something more hedged:

(67)   *Probably*, if it is probably the case that Jill will live if she jumps, then there is a net. (But it might be that there is a trampoline.)

In light of (49) and (67), it is hard to imagine a reason for ruling that embeddings of epistemic vocabulary beyond a certain level of complexity are are semantically uninterpretable. In the absence of such a reason, our theory should deliver semantic values for embeddings of arbitrary complexity. Hence in order to repair the imprecise credence proposal, we would have to model subjects as having not just sets of sets of measures as mental states, but sets of sets of sets of measures, and so on. It is difficult to independently motivate such an arcane model of our mental life.

Second, semanticists like Rothschild must endorse even more complicated models of mental states in order to give a semantics for graded modal vocabulary. It is easy to imagine existential or universal quantification over members of an imaginary mental committee. But graded modals call for probability measures over committee members, and it is difficult to see how one could make sense of this added structure within the imprecise credence model without essentially describing subjects as having precise credences.

The semantics I defend offers a viable alternative in the neighborhood of the imprecise credence proposal. For starters, the semantics extends naturally to graded modals, without requiring that we represent subjects as having mental states more arcane than ordinary credences. As a result, even though it is fairly revisionary to say that contents of sentences are sets of measures instead of sets of worlds, our model of contents can still be defended on the grounds that it simply reflects an independently motivated model of our mental life. In addition, according to our semantics, 'might', 'must', and 'probably' are all type $\langle \mathcal{S}, \mathcal{S} \rangle$, and 'if' is type $\langle \mathcal{S}, \langle \mathcal{S}, \mathcal{S} \rangle \rangle$. Hence complicated sentences like (67) have well-defined semantic values.

Furthermore, our theory even has the resources to say why complicated sentences like (67) might nevertheless sound bad when uttered out of the blue.

The same goes for many sentences containing several referential pronouns. For instance, when uttered out of the blue, (68) sounds questionable at best:

(68)   ?That made that do that to that.

In particular, sentences with several referential pronouns sound bad in isolation when there is a presumption that context will determine different denotations for different pronouns. For instance, (68) sounds worse than (69), just as the nested epistemic vocabulary in (70) sounds worse than the repeated unembedded vocabulary in (71):

(68)   ?That made that do that to that.

(69)   It entered; it saw me; it squealed; and it fainted.

(70)   ?Probably, it is probable that probably Jill will probably live.

(71)   Jill will probably live; John will probably die; Janet will probably cry; and Joe will probably celebrate.

Context often determines different denotations for pronouns in sentences with nested epistemic modals. As a result, a rich context is required for the simultaneous interpretation of the covert pronouns in sentences such as (67) and (70). Here again, in contrast with semantic injunctions against complicated embeddings, pragmatic accounts better fit the contours of our judgments about epistemic vocabulary.

2.4   A semantics for 'probably', 'if', and a covert type-shifting operator

The expression 'probably' has a more complicated semantic function than possibility and necessity modals. The latter modals constrain your credences conditional on propositions in a contextually determined partition. But as a graded modal, 'probably' constrains your credences in members of the partition itself:

$$\llbracket \mathbf{probably}_i \rrbracket^c = [\lambda S.\{m\colon m\left(\bigcup\{p \in g_c(i)\colon m|_p \in S\}\right) > .5\}]$$

In our shorthand: find the union of everyone that accepts that S. If you give that proposition greater than .5 credence, then your credences are contained in the content of 'probably S'.[18] For example, recall that if we are talking about whether Jill is suicidal, you can correctly assert:

18 This semantics follows Kratzer 1991 in taking 'probably' to indicate that something is more likely than not. It is straightforward to adjust the definition so that 'probably' instead indi-

(50)   It is probably the case that Jill will live if she jumps.

The partition relevant to the interpretation of 'probably' in (50) contains two propositions: either Jill is suicidal or she isn't. Just one of these propositions accepts that Jill will live if she jumps, namely the proposition that Jill isn't suicidal.[19] Since you give more than .5 credence to that proposition, your credences are contained in the content of (50), and that is roughly why it sounds okay for you to say it.

At this point, we can also give a more complete explanation of why the content of (40) contains your credences about the outcome of rolling a fair die:

(40)   It is probably$_2$ even or$_1$ less than four.

As mentioned earlier, the sort of context that is hospitable for (40) makes a certain partition salient: either the number rolled is low, or it is high. A second partition is also salient, namely the six possible outcomes of the rolling the die. The first partition determines the content of 'or' and the second determines the content of 'probably'. If you conditionalize your credences on the proposition that the number rolled is low, then you accept that the number is less than four. If you conditionalize your credences on the proposition that the number rolled is high, then you have equal credence in each of the three high number outcomes. Hence you give more than .5 conditional credence to the union of outcomes that accept the number rolled is even. That means your credences conditional on the number being high accept that the number is probably even. It follows from our semantics for 'or' that your credences are in the content of (40), and that is roughly why it sounds okay for you to say it.

Indicative conditionals are semantically like graded modals, insofar as they also constrain your credences in propositions in contextually determined

---

cates likelihood above a contextually defined threshold. In a similar vein, it is straightforward to extend the lexicon of this paper to include other simple epistemic vocabulary, such as 'unlikely', 'at least .3 likely', 'more likely than', and comparative epistemic adjectives.

19 A reminder about our shorthand: your credences satisfy the constraint that a proposition accepts a content just in case your credences conditional on that proposition are contained in that content.

partitions:

$$\llbracket \mathbf{if}_i \rrbracket^c = [\lambda S.\lambda T.\{m\colon m\big(\bigcup\{p \in g_c(i)\colon m|_p \in T\} \\ \big| \bigcup\{p \in g_c(i)\colon m|_p \in S\}\big) = 1\}]$$

In other words, using our shorthand: find the union of everyone that accepts the antecedent of the conditional, and find the union of everyone that accepts the consequent. If you have full credence in the latter proposition conditional on the former, then your credences are contained in the content of the conditional itself.[20]

For example, consider the indicative conditional:

(72)   If$_1$ it is high, it is probably$_2$ even.

The context of (72) makes a certain partition salient: either the number rolled is low, or it is high. The former proposition rejects the antecedent of the conditional, while the latter accepts it. The former proposition also rejects the consequent of the conditional, while the latter accepts it. Hence you have full credence in the union of propositions that accept the consequent of (72), conditional on the union of propositions that accept the antecedent. It follows from our semantics for 'if' that your credences are in the content of (72), and that is roughly why it sounds okay for you to say it.

There is one important respect in which our theory so far is incomplete. I have not yet given a semantics for simple sentences such as:

(65)   Bob is a hire.

(73)   Jill jumps.

(74)   The number rolled is high.

For instance, I have said certain partition propositions "accept that the number rolled is high" or "accept the antecedent of 'if it is high, it is probably even'." This is shorthand for a constraint on probability measures, namely that after conditionalizing on the partition proposition, the resulting measure is contained in the content of (74). Hence simple sentences like (74) must have sets of measures as their contents.

---

20 A disclaimer: this semantics is sufficient to address the motivating concerns of the present paper, but it is not my final word on indicative conditionals. I defend an alternative probabilistic semantics in Moss 2014, motivated by concerns that I have bracketed for ease of exposition here.

There is a natural way of associating simple sentences with sets of measures. According to standard truth conditional semantic theories, the content of a simple sentence is a set of worlds. According to my theory, the content of a simple sentence is the set of measures that assign probability 1 to that set of worlds.[21] This means that the theory need not start from scratch to deliver semantic values for referring expressions, predicates, quantifiers, and so forth. Instead, a covert operator converts traditional semantic values into alternative semantic values:

$$\llbracket \mathcal{C} \rrbracket^c = [\lambda p.\{m \colon m(p) = 1\}]$$

For example, the logical form of the sentence 'Jill jumps' is more accurately represented as '$\mathcal{C}$ Jill jumps'. The semantic value of this sentence is a set of measures, namely $\{m \colon m(\{w \text{ . Jill jumps in } w\}) = 1\}$. Since simple sentences accompanied by the covert operator $\mathcal{C}$ have sets of measures as semantic values, simple sentences can be arguments of type $\langle \mathcal{S}, \mathcal{S} \rangle$ operators and type $\langle \mathcal{S}, \langle \mathcal{S}, \mathcal{S} \rangle \rangle$ operators.

Furthermore, arguments of logical operators can include both simple sentences and sentences containing epistemic vocabulary. For example, the logical form of (40) is more accurately represented as follows:

(40)   [probably$_2$ [$\mathcal{C}$ [it is even]]] or$_1$ $\mathcal{C}$ [it is less than four].

This detail lets us finally give a complete explanation of why your credences are contained in the content of (40) in the context described above. In our most recent explanation of this fact, we said that "if you conditionalize your credences on the proposition that the number rolled is low, then you accept that the number is less than four." The more complete explanation replaces this with the following claim: if you conditionalize your credences on the proposition that the number rolled is low, then *the resulting credence distribution has full credence* that the number is less than four. Fans of gory detail should see Appendix B.1 for an explanation in formal vocabulary.

To sum up so far: I have introduced a semantics for eight expressions, including basic logical operators and epistemic vocabulary. According to this theory, there is a sense in which logical operators *are* epistemic vocabulary. If they occur in the midst of epistemic modals, logical operators deliver

---

21 This content may seem inappropriate, since giving full credence to some proposition is a very strong constraint. In short, I have made some assumptions in order to simplify my discussion, and refinements of the theory in Section 3 address this worry. For a more thorough treatment of these issues, see chapter 2 of Moss 2014.

constraints on credences that depend on what is accepted by propositions in contextually determined partitions. Assigning the same sort of semantic values to logical operators and epistemic vocabulary helps explain the behavior of the latter. The way that 'might' and 'must' and 'probably' interact with each other has a lot in common with the way they interact with logical operators. According to my theory, this is to be expected, as both are interactions between different sorts of epistemic vocabulary.

3 A number of refinements and explanations

I have made three simplifying assumptions in developing the semantics in Section 2. In order to refine the semantics, I will identify these assumptions and say how they can be removed. The first is about the standard effect of assertion, namely that when you hear an assertion with a certain content, you generally come to have credences contained in that content. This claim abstracts away from lying, pretense, supposition, and so on. But more importantly, even in normal cases of assertion, your *credences* do not really come to be contained in asserted contents. The contents of sentences are simply too strong to play that role. The content of a simple sentence is the set of measures that assign probability 1 to some proposition. But it is arguably almost never rational to have full credence in a proposition. Having full credence in a proposition makes you bet on that proposition at arbitrarily risky odds, and makes your belief in that proposition rationally unrevisable by conditionalizing on further evidence. In other words, it makes you have blind faith in a proposition. Assertions rarely if ever have such a dramatic effect.

It might be possible to answer this complaint by saying that our theory governs ideal cases, and that ideal communication really does make subjects have full credence in asserted contents. But even this answer should be accompanied by some suggestions about the effect of assertion in realistic cases. Here is one suggestion: as far as the conversational record is concerned, an act of assertion is a proposal that the content of the assertion be accepted for conversational purposes. For example, suppose you assert that it is raining. Then it will sound bad for either of us to say or even suppose that it might not be raining:

(75)   a.   *Alice:* Oh no. It is raining.
       b.   *Bob:* #If it might not be raining, we should buy some sunglasses.

If your assertion is not challenged or retracted, it does seem that we accept its strong content for conversational purposes. Having accepted that content, Alice and Bob do resemble subjects who would accept bets at arbitrary odds, conversationally speaking, as they cannot even raise the possibility that it is not raining.[22]

In addition to affecting the conversational record, an assertion affects conversational participants. An assertion does not exactly affect your credences, but something more like your *credences for practical purposes.* For example, an assertion of (75a) may have the effect that for practical purposes, it is just as if your credences are contained in its content — that is, when it comes to your preferences and decisions, it is just as if you have full credence in the proposition that it is raining. This account of assertion is designed to mimic contemporary accounts of full belief according to which you believe a proposition when you can treat it as certain for practical purposes. For instance, according to Weatherson 2005, you believe a proposition roughly just in case conditionalizing on that proposition changes none of your preferences over salient options.[23] The analogous account of assertion says you accept an assertion just in case updating your credences on its content changes none of your preferences over salient options. For example, you accept (75a) just in case updating on the proposition that it is raining changes none of your preferences over salient options. In other words, given the analogous account of full belief, you accept (75a) just in case you believe that it is raining. This seems like exactly the right result, as assertions of simple sentences are traditionally taken to constrain your full beliefs. To sum up: given the above accounts of full belief and assertion, you accept an assertion of a simple sentence just in case you believe its content. Even if our accounts of full belief and assertion must ultimately be modified, the latter will deliver intuitive results as long as it mirrors the former.

The second simplifying assumption made in Section 2 is that logical operators have just one semantic value each. In fact, my theory requires a serious and significant revision of this assumption, namely that logical operators have different types of semantic values, depending on whether they embed non-epistemic or epistemic vocabulary. For example, the semantic

---

22 This effect of assertion on the conversational record is elegantly explained by models on which the context set itself is fine-grained. For further discussion, see the context probabilism introduced in Section 8 of Yalcin 2007.

23 Cousins of this principle are defended by Williamson 2000, Ganson 2008, Fantl & McGrath 2010, and Schroeder & Ross 2014.

value of negation given in Section 2 must have a different type of semantic value of negation in simple sentences such as:

(76)   John does not smoke.

For suppose (76) has the following logical form:

(77)   not₁ [𝒞 John smokes]

Then according to the semantics for 'not' in Section 2, the content of (76) contains your credences just in case there is no proposition in the relevant partition such that given that proposition, you have full credence that John smokes. This is not at all what (76) intuitively means. For many partitions, it is very easy for your credences to satisfy this constraint, even if you have a relatively high credence that John smokes. It should intuitively be much harder for your credences to be contained in the content of (76). In fact, in light of our semantics for other sentences without epistemic vocabulary, the content of (76) should intuitively be the set of measures that assign probability 1 to the proposition that John does not smoke.

The appropriate refinement of our semantics involves distinguishing between logical operators that embed epistemic vocabulary and logical operators that embed simple sentences. A simple sentence actually has a set of worlds as its semantic value, which can serve as the argument of a covert type-raising operator. This covert operator need not occur immediately above every simple sentence. In our refined semantics, logical operators can have sets of worlds as arguments. In addition to reinstating traditional semantic values for simple sentences, we reinstate traditional semantic values for logical operators, adding these values to those introduced in Section 2. Hence logical operators have different semantic values in different linguistic contexts: traditional values when their arguments are sets of worlds, and our Section 2 semantic values when their arguments are sets of measures. The logical form of 'John does not smoke' is (78) rather than (76):

(76)   not₁ [𝒞 John smokes]
(78)   𝒞 [not John smokes]

The sentence under the covert operator has a set of worlds as its semantic value, namely the proposition that John does not smoke. Hence the content of (76) is the set of measures that assign probability 1 to that proposition, as desired.

This refinement of our semantics addresses several other potential problems as well. For instance, suppose the logical form of (79) is given by (80):

(79)   John smokes or Jill drinks.

(80)   [𝒞 John smokes] or$_1$ [𝒞 Jill drinks]

Then if the content of (79) contains your credences, there must be some contextually determined partition such that conditional on each proposition in the partition, you either have full credence that John smokes or full credence that Jill drinks. But intuitively you can utter a disjunction even if no such propositions would make you sure of either disjunct. In addition, our semantics should predict that the following inference is valid:

(81)   a. It is not the case that John does not smoke.
        b. Hence: John smokes.

And likewise for the following:

(82)   a. It is not the case that both John smokes and Jill drinks.
        b. Hence: either John does not smoke or Jill does not drink.

However, from the premise that no one accepts that no one accepts that John smokes, we cannot generally infer that John smokes. From the premise that no one accepts that everyone accepts both that John smokes and Jill drinks, we cannot generally infer that everyone either accepts: (a) that no one accepts that John smokes, or (b) that no one accepts that Jill drinks. In other words, if the covert type-raising operator '𝒞' occurs immediately above 'John smokes' and 'Jill drinks' in (81) and (82), the resulting inferences are invalid. Hence our Section 2 semantics does not automatically validate double negation elimination or applications of De Morgan's Laws, even restricted to inferences not containing any epistemic vocabulary.

The above refinement of our semantics validates instances of these inferences where appropriate. For instance, the logical form of 'John smokes or Jill drinks' is given by (83):

(83)   𝒞 [John smokes or Jill drinks]

The semantic value of (83) is the set of measures that assign probability 1 to the proposition that either John smokes or Jill drinks. This semantic value may contain your credences even if no salient information would make you

sure of either disjunct. The logical form of the double negation elimination argument is not (84) but (85):

(84)   a. not_1 not_2 𝒞 John smokes
       b.   Hence: 𝒞 John smokes

(85)   a.   𝒞 not not John smokes
       b.   Hence: 𝒞 John smokes

The logical form of the De Morgan's Law argument is not (86) but (87):

(86)   a. not_1 [𝒞 John smokes and_2 𝒞 Jill drinks]
       b.   Hence: [not_3 𝒞 John smokes] or_4 [not_5 𝒞 Jill drinks]

(87)   a.   𝒞 not [John smokes and Jill drinks]
       b.   Hence: 𝒞 [[not John smokes] or [not Jill drinks]]

It is not hard to verify that the latter inferences are valid, as desired.

   I should emphasize that on the semantics developed here, logical operators are polymorphic. This claim constitutes a loss of theoretical parsimony, which some readers may count as a cost of my theory. Some may even judge that this cost is ultimately too great to be outweighed by the benefits of the theory. However, several facts may help mitigate this cost. For starters, it is a familiar observation that logical operators can embed expressions of various semantic types; indeed, "virtually every major category can be conjoined with 'and' and 'or'" (Partee & Rooth 1983). The theory I defend introduces semantic values with novel semantic types, such as sets of probability measures and functions from sets of measures to sets of measures. In other words, the same sort of semantic type variation in the operators in (88) and (89) is present in (90) and (91):

(88)   John is young, and Mary is young.

(89)   John and Mary are young.

(90)   It is probable that John is young and certain that he is handsome.

(91)   It is probable but not certain that John is young.

A useful and familiar theory of (88) and (89) is that higher-type occurrences of logical operators are the product of type-raising (*cf.* Partee & Rooth 1983 for a canonical early discussion of generalized conjunction and disjunction). This

theory can be extended to (90) and (91) and other uses of logical operators embedding epistemic vocabulary.

In more detail: Partee & Rooth 1982 use a recursive definition to distinguish the *conjoinable* semantic types for which generalized logical operations are defined. Their definition counts both sets of measures, and functions from sets of measures to sets of measures, as conjoinable. Gazdar 1980 proposes simple recursive definitions of generalized conjunction and disjunction, thereby unifying logical operations on different categories. For instance, the function denoted by 'and' always yields either the intersection of its arguments, or the function mapping each element to the intersection of its image under those arguments. If the contextually supplied partition is the trivial partition, then the definitions proposed by Gazdar generate exactly the same semantic values for logical operators as those assigned by my semantic theory. The semantic value of (90) is the intersection of the sets of measures denoted by each conjunct. The semantic value of 'it is probable but not certain' is roughly the function mapping each constraint to the intersection of the set of measures that count its prejacent as probable and the set of measures that count its prejacent as not certain.

In more complicated cases, the relevant type-shifting principles are more complicated than the recursive principles that Gazdar defines. But more complicated type-shifting principles are also not without precedent in the literature; as Partee 1986 notes, type-shifting principles are heterogenous (p. 363). Partee observes that nominalization corresponds to a lexical rule relating the use of 'blue' as an adjective and 'blue' as a proper noun. The situation with higher-type logical operators is similar, as there is a common core of meaning shared by 'and' as a conjunction of propositions and 'and' as a conjunction of sets of measures.

These classic discussions of generalized logical operators provide our theory with useful precedents. They also provide a useful moral, namely that variation in semantic type is not necessarily as costly as unsystematic lexical ambiguity. Partee & Rooth 1983 argue that "the potential disadvantage of having multiple interpretations...is offset by the processing strategy of trying the simplest type first" (p. 13). This argument applies equally well given our semantics for logical operators embedding epistemic vocabulary. The semantic type of a logical operator still uniquely determines its semantic value, just as with the examples discussed by Partee and Rooth. The semantics I defend does not introduce any unforced choices as to what 'and' means in some construction. This fact does not erase the cost of imputing variation in

semantic type to the logical operators, but it should make that cost easier to bear.

A couple of final notes on logical operators. First, we have identified a respect in which 'if' is distinctive among logical operators. Like 'probably' and epistemic possibility and necessity modals, 'if' is thoroughly epistemic. There is no evidence for a second reading of 'if' that accepts sets of worlds rather than sets of measures as arguments. Second, there is also no evidence that logical operators have sets of measures as arguments when they do not embed any overtly epistemic vocabulary. It is not just hard to hear readings riddled with covert type-raising operators, as in (84a) and (86a). They seem genuinely unavailable. In other words, semantic values of sentences are raised from sets of worlds to sets of measures only when forced. If a sentence contains no epistemic vocabulary, then a single type-raising operator scopes over that entire sentence, making it have the right sort of content for assertion. If a sentence does contain epistemic vocabulary, then lower type-raising operators occur only where they are required to make embedded sentences have the right sorts of contents to serve as arguments of that vocabulary. These facts may follow from more general injunctions against unforced type-raising, perhaps along the lines of claims defended in Partee & Rooth 1983, though for reasons of space I shall leave the derivation of these facts as an open question for future investigation.

The third simplifying assumption made in Section 2 concerns the scope of my semantics for epistemic vocabulary. The assumption is that epistemic expressions have just one semantic value each, namely those introduced in Section 2. In fact, sentences with epistemic vocabulary have multiple readings. Just like logical operators, epistemic expressions sometimes have exactly the sort of semantic values that traditional truth conditional semantic theories say they have. In such cases, epistemic expressions do not exhibit any of the behavior that motivates us to reject those semantic theories. For instance, recall that traditional theories come under fire from Yalcin 2007 for failing to predict the infelicity of constructions like the following:

(92)   Suppose that there might be snipers and there are not snipers.

(93)   If there might be snipers and there are not snipers...

However, notice that the following constructions sound fine:

(94)   Suppose that there might — for all you know — be snipers, and there are not snipers.

(95)   If there might — for all you know — be snipers, and there are not snipers…

Furthermore, in the presence of substantial contextual cues, (92) and (93) intuitively mean the same thing as (94) and (95). For instance, imagine that you are in the military, and your instructor gives you the following advice on jungle warfare:[24]

> There are a lot of deadly snipers in the jungle. Before you walk into an area where there are lots of high trees, if there might be snipers hiding in the branches, clear away the foliage with flamethrowers. Do not worry about wasting equipment. Burn the foliage *whenever* there might be snipers. If there might be snipers and there are not snipers, you will have wasted a flamethrower. But if there are snipers and you do not use that flamethrower, you will have wasted human lives.

In the context of this monologue, (93) sounds fine. In short, it sounds equivalent to (95). The arguments in Yalcin 2007 against the standard truth-conditional semantics for 'might' do not succeed in this context.

The same goes for other arguments against the standard semantics for epistemic modals. For instance, imagine that after being trained with the above information, several soldiers enter a jungle warfare situation in which they have the following radio conversation with their commanding officer:

(96)   a. *Soldier:* Should we use flamethrowers to clear the branches?
        b. *Commander:* Is it the case that there might be snipers?
        c. *Soldier:* The scouts have not made a report, so there might be snipers.
        d. *Commander:* Then obviously you should be using your flamethrowers.

Furthermore, imagine that some military students are eavesdropping and judging the jungle soldiers as part of their basic training. The students may say (97), even if they have been informed that there are no snipers in the branches:

---

24 This military monologue is a variation on an example from Egan, Hawthorne & Weatherson 2005. The original example serves a different dialectical purpose, and embeds a simple 'might' sentence rather than a conjunction.

(97)   They should use their flamethrowers, since there might be snipers.

In a similar spirit, the jungle soldiers may later defend themselves by saying that they were right to use flamethrowers, since there might have been snipers. The soldiers may stand by (96c), even if they later find out that there were no snipers in the branches. And finally, in contrast with some other uses of epistemic vocabulary, it is not at all suspiciously hard to identify the correct modal base for 'might' in (96c). Intuitively, the soldiers are simply saying that for all they know, there are snipers in the branches. A traditional contextualist semantics along the lines of Kratzer 1977 seems able to get the content of their utterance exactly right.[25]

It is important to get clear on the dialectical force of such examples. The failure of anti-contextualist arguments when it comes to (96c) does not demonstrate that contextualist semantic theories are sufficient. In the same vein, the success of the same arguments elsewhere does not demonstrate that contextualist theories are unnecessary. The proper reconciliation of our observations is that epistemic expressions have multiple semantic values. In recent literature, facts about embeddings, eavesdropping, and retraction have been understood as supporting anti-contextualist projects. But these facts are better understood as providing us with diagnostic tests. Some uses of epistemic vocabulary exhibit distinctive embedding, eavesdropping, and retraction behavior. Some uses do not. The semantics developed here is a theory of the former, while standard accounts are theories of the latter. In that spirit, facts about eavesdropping should play a role in the literature on epistemic vocabulary like the role played by facts about projection behavior in the literature on presupposition (*cf.* Karttunen 1973, Geurts 1999). In both cases, some distinctive behavior calls out for some modification of a standard theory of content. And in both cases, the behavior itself is so distinctive that it may adequately function as partly constitutive of the sort of language that is best modeled by the modified theory. Just as we have the projection test battery for presuppositional content, we may have a similar test battery for non-propositional content, useful for identifying exactly what uses of epistemic vocabulary the semantic theory in Section 2 is meant to describe.

---

25 The contextualist-friendly uses of epistemic vocabulary include those that are "exocentric," in the terminology introduced by Lasersohn 2005 and adapted to discussions of epistemic modals by Cappelen & Hawthorne 2009 and others. For more examples in a similar spirit, see Section 3 of Dorr & Hawthorne 2012.

In addition to removing three assumptions from Section 2, we may add pragmatic principles to our theory. The Section 2 semantics does not include any principles that distinguish the order in which conjuncts or disjuncts are uttered. Hence supplementary principles must explain phenomena that dynamic semantic theories aim to predict, such as the similarity of the following sentences:

(40)   It is probably even or less than four.

(98)   It's in a funny place, or there's no bathroom in this house.

(99)   John always feels hungry, and he goes to the movies almost every other day.

The alleged facts to be explained are that certain readings of these sentences are unavailable: 'probably' in (40) cannot just be talking about situations where the number rolled is high; 'it' in (98) cannot be referring to the bathroom in the house in question; and 'always' in (99) cannot just be talking about situations where John goes to the movies. Meanwhile, reversing the order of disjuncts or conjuncts makes each of these readings available. According to many dynamic semantic theories, that is because the semantic value of a disjunction depends on the effects of each disjunct on certain local contexts, where reversing the order of the disjuncts changes which local contexts are relevant to those effects. However, pragmatic principles provide strong alternative explanations of the behavior of (98) and (99). Absent any special extra-linguistic context, (100) raises the salience of bathrooms more than 'it's in a funny place' does, and (101) raises the salience of situations in which John goes to the movies more than 'John always feels hungry' does:

(100)   There's no bathroom in this house...

(101)   John goes to the movies almost every other day...

In fact, extra-linguistic context alone is rarely rich enough to determine the referent of 'it' when you say (98). But if you are rather desperately searching around an unfamiliar house while your child is making obvious signs of needing a bathroom, it may sound fine for someone to tell you (98) out of the blue. In the same vein, extra-linguistic context alone is rarely rich enough to determine that 'always' in (99) is just talking about situations where John goes to the movies. But say we have just been debating whether frequent movie-goers become so accustomed to the smell of popcorn that it no longer

makes them feel hungry when they go to the movies. Then the first conjunct of (99) can just be talking about situations where John goes to the movies, exactly as if it had occurred after the second conjunct instead of before.

The Section 2 semantics fits best with a similar pragmatic theory of (40). Talking about rolling a fair die makes a certain partition salient, namely the partition of possible outcomes of rolling the die. It is rare that extra-linguistic context alone is rich enough to determine that some other partition is relevant to the interpretation of 'or' in (40). But it is not impossible, as illustrated by our example of the felicitous utterance of (40). In that example, extra-linguistic context alone makes another partition salient, namely the partition containing the proposition that the number rolled is low and the proposition that it is high. That is why your credences can be contained in the semantic value of (40), as long as you have a high conditional credence that the number is even given that it is high. There are other ways of raising the salience of the same alternative partition, such as using the following sentence fragment:

(102)   The number rolled is less than four, or...

By contrast, the following phrase does not raise the same partition to salience:

(103)   The number rolled is probably even, or...

These facts about salience help explain why it is easier for context to contribute an alternative partition to the interpretation of 'or' in sentences starting with (102) as opposed to (103), but also why it is not impossible for context to contribute an alternative partition in the latter case. It seems inappropriate to promote this explanation to a semantic rule. First, it is difficult to see how a semantic rule could be defeated as necessary in creatively constructed contexts, as pragmatic generalizations are. Second, our pragmatic theory follows from general principles about salience that have little to do with conjunction or disjunction. These concerns by no means settle the debate, but they do support our pragmatic theory as a viable alternative to dynamic semantic accounts of (40).[26]

---

26 Disjunctions like (40) may remind the reader of "modal splitting" cases introduced by Landman 1986. For comparison, it may be helpful to note that Landman does not argue that modal splitting accompanies any disjunction, but merely that sometimes "we have to assume that [some] sentence is added with modal splitting to make sense of it" (p. 205). For a contrasting defense of dynamic semantic theories of disjunction, see Dever 2012. In addition to certain salience facts, using 'either' may enable context to contribute an

A few additional definitions complete our semantic theory. Any context sensitivity enables equivocation in arguments, namely by evaluating earlier claims relative to one context and later claims relative to another. In deciding whether an argument is valid, we must stipulate that we are only concerned about whether its conclusion follows from its premises when all are evaluated relative to a single context. Furthermore, it cannot be that a conclusion follows from some premises just in case any world contained in the latter is contained in the former, since contents are sets of measures rather than sets of worlds. Alternative semantic theories call for an alternative notion of logical consequence. A conclusion follows from some premises just in case every probability measure contained in the latter is contained in the former. Formally, an argument is valid in a context c just in case the intersection of the semantic values of its premises in c is a subset of the semantic value of its conclusion in c. An argument is valid simpliciter just in case it is valid in every context. Finally, it is important to note that the proper objects of semantic evaluation are sentences containing indices. For example, the following sentences constitute one argument:

(104) a. If$_1$ $\mathcal{C}$ low, probably$_2$ $\mathcal{C}$ odd.
       b. Not$_3$ probably$_2$ $\mathcal{C}$ odd.
       c. Hence: not$_3$ $\mathcal{C}$ low.

And the following sentences constitute a distinct argument:

(105) a. If$_1$ $\mathcal{C}$ low, probably$_2$ $\mathcal{C}$ odd.
       b. Not$_1$ probably$_2$ $\mathcal{C}$ odd.
       c. Hence: not$_1$ $\mathcal{C}$ low.

These arguments sound just the same in English. But one may well be valid even if the other is not. This fact accounts for certain behavior of epistemic vocabulary in classically valid arguments.

---

alternative partition to the interpretation of a disjunction, possibly by inducing contrastive focus (*cf.* Hendriks 2004). For reasons of space I shall leave this proposal as an open question for future investigation.

4 How our theory accounts for the behavior of epistemic vocabulary

Now for the payoff. The semantics in Section 2 and pragmatics in Section 3 account for the distinctive behavior of epistemic vocabulary described in Section 1.

4.1 Nested epistemic vocabulary

The semantics in Section 2 explains why nested epistemic modals signal that different opinions about some subject are in play. A sentence with nested epistemic modals constrains your credences conditional on contextually determined propositions. These conditional credences are different opinions about some subject. For example, remember that you might say (64) if you are torn between various ways of evaluating job candidates:

(64)   It is definitely$_1$ the case that Bob might$_2$ be the best candidate for the
         job.

As mentioned in Section 2.2, on the most natural reading of (64) in this context, the partition used to interpret 'definitely' contains propositions about what virtues matter when evaluating candidates. Just settling what virtues matter does not leave you entirely certain which candidates are best for the job, since you may still be unsure which candidates have which virtues. The different opinions that are salient in the context of (64) are credences conditional on propositions about what sorts of virtues matter. The content of (64) is the constraint that each opinion consider it possible that Bob is the best candidate for the job, conditional on any proposition about which candidates have which virtues.

It is easier to make sense of nested epistemic modals by imagining that each conditional credence distribution is the opinion of some expert, since it is easier to make sense of internal facts by imagining that they correspond with external ones. If some experts most value teaching experience and others most value research quality, for instance, conditional credences governed by (64) may simply match unconditional credences that actual experts have. The subject forms her credences by splitting the difference between expert credences, weighting the credence of each expert according to her credence that they are trustworthy with respect to what sorts of virtues matter when evaluating candidates. In cases where there are no obvious facts about which imagined experts disagree — such as what sorts of virtues matter in evaluating

candidates — it is harder to associate each expert with an element of some obvious partition associated with the outer modal. But it does seem that whenever we imagine a subject splitting the difference between various expert credences, we still visualize that compromise as if there are various propositions that she splits her credence between. In other words, we imagine the subject acting as if there is a sense of 'trustworthy' according to which she should simply defer to the most trustworthy expert, and then simply splitting her credence between propositions about which expert is most trustworthy. These propositions form an artificial partition, and we imagine that partition being contributed by context to the interpretation of the outer modal.

The semantics in Section 2 can also explain the behavior of contradictory nested epistemic modals. For example: recall that you could correctly utter (60) after a fair die has been placed under a cup according to whether the number rolled was low or high:

(60)   It *might* be that the number rolled is probably even.

In the same circumstances, since you have just .5 credence that the die landed on an even number, you could also correctly utter (106):

(106)   It is not probably even.

But you could not utter the conjunction of these sentences:

(107) #The number rolled might be probably even, and it is not probably even.

It sounds fine to utter (60) and (106) separately because 'might' and 'not' are most naturally interpreted using different partitions. It is most natural to say (60) when you are thinking about which cup the die is under. The modal 'might' is interpreted using the partition that the number rolled is either low or high. (60) means that someone in that partition accepts that the number rolled is probably even, and your credences satisfy that constraint. It is most natural to say (106) when you are not most concerned with which cup the die is under, but rather with your all-things-considered credences about the die. The operator 'not' is interpreted using the trivial partition. (106) means that no one in that partition accepts that the number rolled is probably even, and your credences also satisfy that constraint.

This derivation demonstrates some nice features of the Section 2 semantics. First, our intuitive feeling that someone is talking about all-things-

considered credences corresponds to a specific feature of semantic interpretation, namely that context contributes a trivial partition to the interpretation of an epistemic modal. Second, we can say why the content of epistemic vocabulary under negation often seems to be the complement of the content of that epistemic vocabulary. In the absence of defeating contextual cues, subjects commonly use epistemic expressions under negation to constrain all-things-considered credences. Hence 'not' is interpreted using the trivial partition, thereby denoting the simple operation of taking the complement of the content of the embedded sentence.

If you utter (60) and (106) together, by contrast, it is natural to interpret 'might' and 'not' using the same partition. This follows from more general facts about the interpretation of contextual parameters. For instance, it is natural to hear the following sentence as expressing a contradiction:

(108) #No one danced and everyone danced.

It is hard to hear 'no one' and 'everyone' as quantifying over different domains in (108), even if you are inclined to be charitable. For similar reasons, it is hard to interpret 'might' and 'not' using different partitions in (107). The same goes for our original conjunction:

(24)   #It is not the case that it is probably John and it might be the case that it is probably John.

This sentence sounds bad when some experts believe that a certain masked murderer is probably John, but most experts believe it is probably Mary. In isolation, the first conjunct expresses the constraint that you side with the majority in forming your all-things-considered credences. The second expresses the constraint that you not ignore the possibility that the minority opinion is most trustworthy. But when the conjuncts are put together, it is hard to interpret them using such different partitions.

Finally, our theory can explain the intuitive connection between nested modals, evidential weight, and credal resilience. Recall that Eric can say (27) instead of (25) because his high credence that Liem is wearing green is resilient, and based on a lot of evidence:

(25)   It might be the case that Liem is probably wearing green.

(27)   It is almost certainly the case that Liem is probably wearing green.

Joyce 2005 suggests that evidential weight manifests itself in credal resilience *when your credences are mediated by chance hypotheses.* In particular, "the weight of evidence tends to stabilize [your] credence in a particular way: it stabilizes credences of chance hypotheses, while concentrating most of the credence on a small set of these hypotheses" (p. 166). Eric's credences about what Liem is wearing are mediated by propositions relevantly like chance hypotheses, namely claims about how much Liem likes wearing green. The more Liem likes wearing green, the higher the objective chance that he will wear green on any given day. Having seen Liem wear green on 500 out of 800 days, Eric concentrates almost all of his credence on specific hypotheses about exactly how much Liem likes wearing green. That is why the weight of Eric's evidence makes his high credence that Liem is wearing green so resilient.

That is also why Eric can say (27). The sentence (27) is intuitively interpreted using just these same hypotheses about how much Liem likes wearing green. In shorthand: Eric gives almost all his credence to hypotheses that accept that Liem is probably wearing green. According to our semantics, that is why his credences are contained in the content of (27). Madeleine has much less evidence. She may give a majority of her credence to hypotheses that accept that Liem is probably wearing green. But she should still give considerable credence to a number of other hypotheses compatible with his wearing green on 5 out of 8 days. For instance, it is compatible with Madeleine's evidence that Liem likes red more than green, but dumped all his red shirts in the laundry just before the 8 days in question. Hence according to our semantics, her credences are not contained in the content of (27). Having a lot of evidence simultaneously makes Eric's high credence that Liem is wearing green very resilient, and licenses his embedding 'Liem is probably wearing green' under very strong epistemic modals. In more generality: when your credences are mediated by something like chance hypotheses, increasing evidential weight simultaneously makes your credences more resilient and licenses your embedding corresponding constraints under stronger epistemic modals.

4.2   Epistemic vocabulary under disjunction

The theory in Section 2 and Section 3 explains why you can assert (38) even when you cannot assert its first disjunct and you can deny its second disjunct:

(38)   It is less than four or probably even.

As discussed in Section 3, the first disjunct of (38) makes a certain partition salient, namely that the number rolled is either low or high. In the resulting context, (38) has just the same content as (40):

(40)   It is probably even or less than four.

As discussed in Section 2.4 and Appendix B.1, your credences are contained in that content. That is why you can assert (38). This holds even though your credences are not contained in the content of the first disjunct, since you do not have full credence in the proposition that the number rolled is less than four. And it holds even though your credences are contained in the content of the negation of the second disjunct, for reasons just reviewed in Section 4.1. This discussion highlights another nice feature of our semantics: your credences may be contained in the content of a disjunction even when they are not contained in the content of either disjunct, as long as all members of the salient partition accept at least one of the disjunct constraints. This feature alleviates the central concern about 'or' exportation that Schroeder 2012 raises for the semantics in Yalcin 2007.

In the meantime, disjunctive syllogism remains valid for disjunctions free of epistemic vocabulary. For example, consider the disjunction:

(79) $\mathcal{C}$ [John smokes or Jill drinks]

If the content of (79) contains your credences, then you have full credence that either John smokes or Jill drinks. If you can also deny the second disjunct, you have full credence that Jill does not drink. From this it follows that you have full credence that John smokes, and hence that you can assert the first disjunct by itself.

In addition, our Section 3 theory explains why disjunctions such as (40) sound felicitous in some contexts and infelicitous in others. In some contexts, the partition that the number rolled is low or high is salient. In those contexts, 'or' in (40) is interpreted using that partition, and the disjunction sounds fine. In some contexts, the partition that the number rolled is low or high is not salient. In those contexts, 'or' is interpreted using the trivial partition, and the disjunction sounds bad. In short, absent contextual cues, it is hard to know what non-trivial partition you might have in mind when you say the first half of (40), just as it is hard to know what specific situations you might have in mind when you say the first half of (99):

(99)   John always feels hungry, and he goes to the movies almost every other day.

To sum up: disjunct order makes a purely pragmatic contribution to the content of a disjunction. If extra-linguistic context does not make a certain partition salient, using disjuncts in a certain order can have that effect. But if extra-linguistic context already makes that partition salient, changing around the order of the disjuncts has no additional semantic effect. That is why reversing disjunct order does not affect the interpretation of disjunctions like (40) in contexts where they are already felicitous.

4.3   Epistemic vocabulary over indicatives

The semantic theory given in Section 2 explains why you can assert (44) when we are talking about whether there is a net along the roof of your office building, but not when we are talking about whether Jill is suicidal:

(44)   Probably, if Jill jumps off the building, she will die.

The logical form of (44) is as follows:

(109)   probably$_1$ [if$_2$ [$\mathcal{C}$ Jill jumps off the building]] [$\mathcal{C}$ she will die]]

In shorthand: the content of (44) contains your credences just in case you give more than .5 credence to the union of everyone in $g_c(1)$ who accepts that Jill will die if she jumps. Say we are talking about whether there is a net along the roof. Then $g_c(1)$ contains two propositions: that there is a net, and that there is no net. The latter proposition accepts that Jill will die if she jumps. And you give that proposition more than .5 credence. Hence according to our semantics, the content of (44) contains your credences. That is why you can assert (44).[27] By contrast, say we are talking about whether Jill is suicidal. Then $g_c(1)$ contains different propositions: that Jill is not suicidal, and that she is suicidal. The latter proposition accepts that Jill will die if she jumps. But you do not give that proposition more than .5 credence. And the former proposition does not accept that Jill will die if she jumps. Hence according to our semantics, the content of (44) does not contain your credences. That is why you cannot assert (44).

---

27 Similar reasoning explains the observation by Kaufmann 2004 that when certain partitions are salient, the credence that we assign to 'if I pick a red ball, it will have a black spot' fails to match our conditional credence in the consequent given the antecedent.

The semantic theory given in Section 2 also explains why you can correctly assert (49) when we are talking about whether there is a net along the roof:

(49)   If$_1$ it is probably$_2$ the case that Jill will live if$_3$ she jumps, then there is a net.

In shorthand: the content of (49) contains your credences just in case you have full conditional credence in the union of everyone in $g_c(1)$ who accepts that there is a net, given the union of everyone in $g_c(1)$ who accepts that it is probably the case that Jill will live if she jumps. Again, if we are talking about whether there is a net, $g_c(1)$ contains two propositions: that there is a net, and that there is no net. The former proposition accepts that it is probably the case that Jill will live if she jumps. The latter does not. Since you have full conditional credence that there is a net, given that there is a net, your credences are indeed contained in the content of (49). That is why you can assert it when we are talking about whether there is a net along the roof.

The context dependence of (49) and (50) accounts for why you can assert either premise separately but you cannot conclude (51):

(49)   If it is probably the case that Jill will live if she jumps, then there is a net.

(50)   It is probably the case that Jill will live if she jumps.

(51)   There is a net.

The content of each premise depends on whether we are talking about whether there is a net along the roof, or about whether Jill is suicidal. Your credences are contained in the content of (49) in the former but not the latter contexts. And they are contained in the content of (50) in the latter but not the former contexts. There is no single context in which both premises are assertable, and hence no context in which you are licensed in concluding (51) on the basis of (49) and (50).

4.4   Epistemic vocabulary in classically valid arguments

Recall from Section 1.5 that the following argument seems invalid when we are reasoning about the outcome of rolling a fair die:

(52)   a.   If it is low, it is probably odd.
       b.   It is not probably odd.
       c.   Hence: it is not low.

As discussed in Section 4.1, sentences like (52b) intuitively concern your all-things-considered credences, and so the negation operator in (52b) is naturally interpreted using the trivial partition. As discussed in Section 4.2, linguistic context can make non-trivial partitions salient. For instance, recall that 'it is high or…' makes a certain partition salient, namely that the number rolled is either low or high. The construction 'if it is low…' makes the same partition salient, and so the conditional operator in (52a) is naturally interpreted using this partition. Hence the most natural reading of (52) has the following logical form:

(110) a.   if$_1$ [$\mathcal{C}$ it is low] [probably$_2$ $\mathcal{C}$ it is odd]
       b.   not$_3$ [probably$_2$ $\mathcal{C}$ it is odd]
       c.   Hence: $\mathcal{C}$ not [it is low]

This argument is indeed invalid. There are contexts in which the semantic value of its conclusion does not contain the disjunction of the semantic values of its premises. See Appendix B.2 for a formal proof.[28]

A similar diagnosis follows for our troublesome instance of constructive dilemma:

(53)   a.   If it is low, it is probably odd.
       b.   If it is high, it is probably even.
       c.   It is either low or high.
       d.   Hence: either it is probably odd or probably even.

---

28 As an editor helpfully points out, there seems to be a special problem with *modus tollens* that there isn't for *modus ponens*: it is easier to hear invalid interpretations of the former. This felt difference between *modus ponens* and *modus tollens* corresponds to a theoretical difference between these inferences. In particular, the conclusion of a *modus tollens* inference such as (110c) contains a negation operator embedding a simple sentence. Standard injunctions against unforced type-raising predict that this occurrence of 'not' will express propositional negation. Meanwhile, the constraint embedded in the second premise of (110) forces a higher-type interpretation of 'not' in that premise. Hence the determination of the semantic type of 'not' in *modus tollens* inferences ensures that these inferences are given invalid interpretations, whereas no similar mechanism leads us to hear *modus ponens* inferences as invalid.

The conclusion of this argument sounds bad because 'or' in (53d) is naturally interpreted using the trivial partition, and you neither have greater than .5 credence that the number rolled is even, nor greater than .5 credence that it is odd. In other words, the most natural reading of (53) has the following logical form:

(111) a. if$_1$ [$\mathcal{C}$ it is low] [probably$_2$ $\mathcal{C}$ it is odd]
       b. if$_1$ [$\mathcal{C}$ it is high] [probably$_2$ $\mathcal{C}$ it is even]
       c. $\mathcal{C}$ [it is low or it is high]
       d. Hence: [probably$_2$ $\mathcal{C}$ it is odd] or$_3$ [probably$_2$ $\mathcal{C}$ it is even]

This argument is invalid. See Appendix B.3 for a formal proof.

These inferences share a couple of features that are responsible for their invalidity. First of all, the major logical operators in these arguments are not co-indexed. A non-trivial partition is used to interpret 'if' in the first premise of (52), while a trivial partition is used to interpret 'not' in the second. A non-trivial partition is used to interpret 'if' in the first two premises of (53), while a trivial partition is used to interpret 'or' in the conclusion. In other words, there is an equivocation in these arguments. The arguments are just as bad as arguments where context naturally supplies different domains of quantification to overt quantifiers, such as:

(112) a. Everyone failed the exam.
       b. Everyone who failed the exam took the exam.
       c. Hence: everyone took the exam.

Second of all, certain logical operators have different semantic values in the premises and conclusions of these arguments. The operator 'not' has higher type in the second premise of (52) and lower type in its conclusion. The operator 'or' has lower type in the third premise of (53) and higher type in its conclusion.

Although it is not directly relevant to our discussion of (52) and (53), it is worth mentioning a third feature that may be responsible for the invalidity of arguments containing epistemic vocabulary. I have occasionally assumed that probability measures assign some probability to each member of contextually determined partitions. This simplifying assumption deserves much more discussion than I can give it in this paper. A number of arguments become more complicated once we allow that contextually determined partitions may contain propositions in which you have absolutely no credence. For example,

consider the following alleged counterexample to *modus ponens* from McGee 1985:

(113)   a.   If a Republican wins the election, then if it's not Reagan who wins it will be Anderson.
        b.   A Republican will win the election.
        c.   Hence: if it's not Reagan who wins, it will be Anderson.

It may be that (113b) is assertable because for practical purposes, you have full credence that Reagan will win the election. Then according to my theory, whether you can assert (113c) will depend partly on your credences conditional on some proposition in which you have absolutely no credence, practically speaking. This complicates our evaluation of the validity of (113).

If we restrict our attention to arguments without these troublesome features, we can resurrect classical inference rules. Recall from Section 3 that sentences containing indices are the proper objects of semantic evaluation. Hence strictly speaking, arguments containing indices are evaluated for validity. There is an entire family of *modus tollens* arguments. One such argument is invalid. But our acceptance of other instances of *modus tollens* is justified. The same goes for constructive dilemma arguments. In more precise terms: let us say that a context c is *well-behaved* with respect to a measure m and an argument A just in case for every index i on epistemic vocabulary in A, we have $m(p) > 0$ for all propositions $p \in g_c(i)$. Let us say that an argument A is *quasi-valid* just in case for every measure m and every context c that is well-behaved with respect to m and A, if m is contained in the semantic values of the premises of A in c, then m is contained in the semantic value of the conclusion of A in c. Then the following argument schema is quasi-valid:

(114)   a.   if$_1$ P, Q
        b.   not$_1$ Q
        c.   Hence: not$_1$ P

In other words, replacing each letter in the above schema with a sentence whose semantic value is a set of measures always yields a quasi-valid argument. The same goes for the following argument schema:

(115)   a.   if$_1$ P, Q
        b.   if$_1$ R, S
        c. P or$_1$ R
        d.   Hence: Q or$_1$ S

In addition, instances of chancy *modus ponens* are valid simpliciter:

(116)   a.   probably$_1$ P
        b.   if$_1$ P, Q
        c.   Hence: probably$_1$ Q

Thus our semantics vindicates our attachment to certain instances of *modus tollens* and constructive dilemma, even instances riddled with epistemic vocabulary (*cf.* Appendices B.4 and B.5 for proofs). In addition, our semantics validates any instance of chancy *modus ponens* (*cf.* Appendix B.6). As long as we are in a well-behaved context, we are perfectly justified in inferring according to the argument forms given above.

I noted above that 'or' in the conclusion of (53) is interpreted relative to the trivial partition, and that this helped explain why the inference was invalid. It is instructive to note that on my theory, 'or' in the conclusion of a constructive dilemma may not always be interpreted relative to the trivial partition. In fact, sometimes context supplies a non-trivial partition and informants actually accept the conclusions of constructive dilemma inferences similar to (53).

This phenomenon is already familiar in the case of constructive dilemma inferences embedding deontic vocabulary. For example, suppose that a bunch of miners are trapped together in one of several shafts. We can either save all the miners by blocking the correct shaft or kill just one miner by blocking no shaft. Consider the following inference:

(117)   a.   If the miners are in shaft A, we ought to block A.
        b.   If the miners are in shaft B, we ought to block B.
        c.   The miners are in shaft A or shaft B.
        d.   So either we ought to block A or we ought to block B — we just don't know which!

There is an acceptable reading of (117d) in this inference, on which it communicates roughly that certain further evidence would either decisively recommend that we block A or decisively recommend that we block B.

The same phenomenon is less familiar, but just as present with constructive dilemma inferences embedding epistemic vocabulary. For example, suppose that you have taken a test to determine how likely it is that you have cancer, where the test produced a colored line that is either red or green or blue. Red means that cancer is very likely; green means that it is very unlikely; and blue means that the test failed to yield any useful information.[29] The doctor tells you that the test did not produce a blue line. Then you may infer:

(118)   a.   If the test line was red, it is very likely that I have cancer.
         b.   If the test line was green, it is very unlikely that I have cancer.
         c.   The test line was red or green.
         d.   So either it is very likely that I have cancer or very unlikely that I have cancer — we just don't know which!

Just as with (117d), many informants hear an acceptable reading of the conclusion (118d). On this reading, (118d) communicates roughly that certain further evidence would either strongly confirm that you have cancer or strongly disconfirm that you have cancer. Happily, my theory naturally accommodates this reading. The 'or' in (118d) may be interpreted relative to a contextually salient partition, namely the possible test results. This reading of (118d) will contain your credences after you update them on (118c), since each remaining test result will confirm one of the disjuncts of (118d). Hence (118d) will sound acceptable in this context.

Of course, we cannot conclude that context will always supply some non-trivial partition whenever an instance of constructive dilemma is uttered. In fact, if we remove 'we just don't know which' from (118d) and instead say that red and green are equally likely test results, informants may start hearing the conclusion as unacceptable. The point here is just that my theory accommodates some contextual variability in whether we hear constructive dilemma inferences as valid or invalid, and that this is a benefit rather than a cost of my theory.

4.5   The scope of my theory and avenues of further research

In assessing the explanations in Sections 4.1–4.4, readers may observe that my semantic theory yields predictions about concrete cases only in conjunction with supplementary assumptions to the effect that certain contexts

---

29 I am grateful to Josh Dever for suggesting that I discuss this sort of example here.

make certain partitions available as the values of certain pronouns. This may cause some readers to worry that the theory itself does not generate strong enough predictions to constitute an interesting piece of research. In response, it is worth noting that it would be unusual to demand that a contextualist theory come equipped with conditions for deriving the available values of covert pronouns from facts about the context. To compare a salient example, compare the defense of contextualism in Kratzer 1981. Kratzer does not equip her readers with hard and fast rules for determining what accessibility relations could constitute the modal base supplied by a particular context. It is not hard to understand why. The facts about context that determine the availability of pronoun values are so subtle that it may be literally impossible for any theorist to provide rules that yield predictions about arbitrary examples that readers may construct themselves.

Even in the absence of such rules, however, it is possible to give strong arguments in support of a contextualist theory of some vocabulary. In short, it is possible to give "non-constructive proofs" that the semantic value of an expression depends on certain sorts of contextual parameters. These arguments do not establish that my theory correctly predicts that a particular reading of a sentence is available in a particular context. Instead they present indirect evidence that we use epistemic vocabulary to express different constraints on our conditional credences in different contexts.

The indirect arguments supporting my theory highlight five features of sentences containing epistemic vocabulary. In brief: facts about partitions affect whether sentences containing epistemic vocabulary are acceptable in various contexts. In some contexts, sentences containing epistemic vocabulary express multiple constraints on conditional credences. There are infelicitous reports of disagreement between speakers who utter apparently contradictory sentences in contexts where different partitions are salient. The behavior of contextually supplied partitions mimics attested behavior of contextually supplied quantifier domains. Finally, sentences containing epistemic vocabulary exhibit binding effects where different partitions are relevant to the interpretation of different values of the bound variable.

The first three of these five features are closely related. For starters, recall that facts about salient questions affect whether (44) or (45) is assertable in a context:

(44)   Probably, if Jill jumps off the building, she will die.

(45)   Probably, if Jill jumps off the building, she will live.

(44) is assertable when someone is wondering whether there is a net along the roof of the building, and (45) is assertable when someone is wondering whether Jill is suicidal. This means that in contexts where it is not clear whether the first or second question is at issue, sentences like (44) can have multiple readings. For instance, you might say (120) to clarify which reading of (119) is intended in a particular context:

(119)   Is it likely that Jill will die if she jumps off this building?

(120)   Wait a minute. Are you just asking whether I think there is still a safety net along the roof, or are you asking whether I think Jill may be ready to jump from the roof even though — I am pretty sure — the net isn't there anymore?

In the absence of clarification, though, you may use (44) or (45) to express consistent attitudes about Jill, namely because you may consistently believe both that there is no net along the roof and that Jill would not jump unless there was a net. That is also why it can sound bad to report someone who says (44) and someone who says (45) as disagreeing about likelihood facts. They may have consistent beliefs and differ only with respect to which beliefs they would use (44) to report. In the same vein, insofar as you are not sure which answer to give to (119), it is not that you cannot make up your mind about whether Jill will probably live if she jumps. In both cases, you are perfectly clear about the facts, and merely unclear about which facts you would be using the conditional to report.

In addition to indicative conditionals, graded modal vocabulary may have different readings in different contexts. For example, suppose Alice, Bob, Casey, and Dylan are among your good friends. Alice and Bob are married, and Casey and Dylan are married. There is a party next door, and you are wondering whether your various friends are there. Alice and Casey are best friends with each other and often go to parties together. The same goes for Bob and Dylan. But none of them really enjoys going out with their best friend's spouse, so all four friends seldom end up at the same party together. In a normal context, your knowledge of these facts is enough to license your saying:

(121)   If Alice is more likely than Bob to be at the party, then Casey is more likely than Dylan to be there.

If you were to find out that Alice was almost certainly at the party, for instance, then you would conclude that Casey was almost certainly there, while Dylan was almost certainly absent. But here is a twist: suppose that someone overhears you say (121) and replies:

(122) But what if the reason that Alice is more likely than Bob to be at the party is that Bob has just recently been diagnosed with some serious illness? Then Casey and Dylan would both surely be present to provide Alice with emotional support.

After someone says (122), it is hard to continue affirming (121). The speaker of (122) raises the possibility that Alice is more likely than Bob to be at the party in virtue of Bob having some serious illness. Conditional on that possibility, you accept the antecedent of (121), reject its consequent, and reject the conditional itself. Hence you no longer take yourself to be in a position to say (121). But at the same time, you may stand by your original reason for saying it. If you had to update your credences on the information that Alice is more likely than Bob to be at the party, your resulting credence that Casey was there would be higher than your credence that Dylan was there.[30] In fact, you may even feel that it is somewhat unfair to criticize your original assertion by introducing an obscure possibility into the conversation. The more obscure the possibility, the more you may feel as if your interlocutor is changing the subject rather than taking what you said at face value, though it may still be hard to ignore possibilities once they have been introduced.

This discussion should sound familiar, namely because natural language quantifiers exhibit just the same behavior. In a normal context, you might say (123) only to have someone reply with (124):

(123) There is nothing in the fridge.

(124) But what about condiments? Isn't there mustard in the refrigerator door?

---

30 In this example, (122) raises a possibility that is probabilistic in nature. The possibility in the antecedent of (121) is also probabilistic; if you update on that possibility, your credences are updated on a non-propositional constraint. This sort of learning is governed by generalizations of classic updating rules. See Diaconis & Zabell 1982 for an introduction to this method of maximum entropy updating, and see Yalcin 2007 and Moss 2013 for discussion of this sort of updating in the context of semantic theories of epistemic vocabulary.

After someone says (124), it is hard to continue affirming (123), even though you may stand by your original reason for saying it. In fact, you may even feel that it is unfair to criticize your original assertion by introducing unanticipated refrigerator contents into the conversation. The less anticipated the contents, the more you may feel as if your interlocutor is changing the subject; it feels inappropriate to challenge (123) on the basis of dust or air molecules, for instance, though it may still be hard to contract the domain of quantification once it has been expanded.

It is easy for my theory to explain these similarities between epistemic vocabulary and quantifiers, namely because on my theory, introducing potential probabilistic evidence just amounts to expanding a contextually supplied domain of quantification. The informal gloss of my semantics explicitly uses 'everyone' and other quantifiers to highlight this feature of the semantic theory. (122) and similar sentences expand the domain over which 'everyone' quantifies.

For example: when you utter (121), you may be considering several simple likelihood assessments that you might end up with. For instance, you may get evidence that Alice is more likely or equally likely or less likely to be at the party than Bob. The same likelihood assessments will say that Casey is more likely or equally likely or less likely to be at the party than Dylan, respectively. The indicative (121) is accepted by each member of this contextually supplied partition, and that is why it sounds just fine. By contrast, uttering (122) introduces more complicated likelihood assessments into the contextually supplied partition, including your credences conditional on the information that Alice is more likely to be at the party than Bob for reasons related to serious illness. (121) is not accepted by each member of this expanded partition, and that is why it no longer sounds fine.

In addition to these asymmetries in context shifting, epistemic vocabulary shares another feature with paradigmatic implicit arguments, namely that both exhibit binding effects. For example, many accept that a covert location argument in (125) is bound in sentences such as (126):

(125)   The local bar is closed.

(126)   Everywhere John goes, the local bar is closed.

In just the same way, quantifiers can bind covert partition arguments of sentences like (127), namely in sentences such as (128):

(127)   Alice might be a probable hire.

(128)   However you look at it, Alice might be a probable hire.

Suppose that Alice is a candidate for a job in our department. There are two ways of collecting information about the likelihood that we will hire her. First, Alice has several colleagues, each of whom has expressed an informal opinion about whether she is a probable hire. Second, we have collected several formal letters, each of which contains an opinion about whether she is a probable hire. At least one colleague says that she is a probable hire, and at least one letter says that she is a probable hire. In this context, even if we are unsure whether informal or formal opinions are more trustworthy, we may still accept (128). To spell out this reading: we accept (128) because regardless of whether we form our credences about Alice by collating informal or formal likelihood assessments, we give some credence to an assessment according to which she is a probable hire. This reading is easy for my theory to explain, since on my theory, 'however you look at it' can bind the covert partition argument of 'might' in (128).

   This behavior of 'however you look at it' is not without precedent. For instance, suppose that causation is essentially contrastive, and that 'cause' has a covert partition argument.[31] If the Indian government and a horrible drought are each causally relevant to a famine, then a contextually supplied partition of alternatives may determine whether (129) or (130) is acceptable:

(129)   The Indian government caused the famine.

(130)   The drought caused the famine.

But now suppose that the Indian government was secretly depleting ground-water resources in a way that caused the drought. Then the following will sound acceptable:

(131)   However you look at it, the Indian government caused the famine.

There are multiple sets of causal alternatives relevant to (129), including sets of theories about potential political causes and sets of theories about potential meteorological causes. (131) is acceptable because any such set of alternatives will contain some theory that says that the Indian government is responsible for the famine.

---

31 See Mackie 1974 for a classic discussion of causal contrastivism and Schaffer 2005 for a more detailed discussion of the semantics of causal claims. The famine example is due to Hart & Honoré 1985.

It is just the same with (128): there are multiple sets of theories relevant to our judgment that Alice is a probable hire, including sets of informal opinions and sets of formal opinions. (128) is acceptable because any such set will contain some theory that says that Alice is a probable hire. To sum up: binding effects in sentences like (128) are evidence for a contextualist theory of epistemic vocabulary, where the bound argument of the higher modal ranges over various ways of generating likelihood assessments — that is, various partitions of propositions about what evidence should be trusted in assessing the embedded probabilistic claim.

The above arguments constitute indirect evidence for my theory. It is also possible to argue for the theory more directly, namely by assessing substantive predictions delivered by my theory in conjunction with supplementary principles about available values of partition pronouns. For starters, my theory entails that for many cases of nested epistemic modals, if context supplies both the inner and outer modal with the same simple partition, the resulting semantic value could be expressed using just one modal. For example:

(132)   It is unlikely that Alice is a probable hire.

Suppose we evaluate both 'unlikely' and 'probable' using the partition consisting of the claim that Alice is a hire, and the claim that she is not. According to my theory, the result is the set of measures that give less than .5 credence to the claim that Alice is a hire. According to my theory, that set of measures is also the semantic value of the sentence:

(133)   It is unlikely that Alice is a hire.

Hence general pragmatic reasoning should lead us to interpret at least one modal in (132) relative to a fine-grained partition. Against the background of Gricean advice against unnecessary prolixity, we should charitably interpret (132) as having some meaning that could not have been more succinctly expressed using (133). The same goes for many nested modal constructions. To sum up: my theory predicts that nested modals are naturally interpreted as being about higher-order information.

A number of other useful results follow from facts about what questions are commonly addressed by certain sentences. For instance, modals embedding a simple prejacent are commonly used to address the question of whether or not that prejacent is the case. For example, the following sen-

tences are commonly used to address the simple question of whether John is guilty:

(134)   John is probably guilty.

(135)   John might be guilty.

(136)   If John is guilty, then…

This fact helps explain the result that 'probably' is commonly used to express the simple constraint that you give more than .5 credence to its prejacent. Consider any context where the proposition that John is guilty is a member of the contextually supplied partition, or even just a member of the algebra generated by the contextually supplied partition elements. In any such context, my semantics for 'probably' entails that you accept (134) just in case you give more than .5 credence to the proposition that John is guilty.[32]

By contrast, sometimes (134) is instead used to directly address more complicated probabilistic questions such as:

(137)   Is it probable that John is guilty?

For example, suppose that you are sitting on a jury and you hear several experts testify about the likelihood that John committed a certain murder. A large majority of the experts say that John is probably innocent, but some experts insist that John is probably guilty. Suppose the judge then calls you to answer (137). There are multiple reactions you may have. On the one hand, you may feel as if you do not know the answer to the question, since you are not in a position to rule out the expert testimony of the minority. But on the other hand, you may feel as if you should simply answer 'no' to the question, since you have greater than .5 credence that the large majority of experts have formed the right opinion about John. This multiplicity of readings of (137) reflects a multiplicity of underlying questions that the judge could be addressing. On the one hand, the judge may be interested in whether she should strike the testimony of particular experts as misleading. On the other hand, she may simply be wondering whether she should give more than .5 credence to the first-order claim that John is guilty. In the first-order context, the judge uses (137) to address the question of whether John is guilty, and an affirmative answer expresses a simple constraint. By

---

32 This result follows from the axiom of finite additivity, since for the purposes of this paper, I am setting aside cases where the contextually supplied partition has infinitely many elements.

contrast, in situations where speakers are explicitly assessing the reliability of various pieces of evidence, 'probably' sentences are used to express more complicated constraints on credences.

The same sort of generalization helps explain the behavior of simple 'might' sentences. For example, 'John might be guilty' commonly raises the question of whether John is guilty. Hence 'John might be guilty' not only relies on the contextually supplied partition for its semantic value, but affects that same partition. And as long as the algebra generated by the partition contains the claim that John is guilty, my theory predicts that 'John might be guilty' will sound acceptable to any subject who has some credence that John is guilty, since that subject must therefore end up giving some credence to at least one partition element that entails that John is guilty.[33]

Finally, the same sort of generalization applies to antecedents of indicative conditionals. For example, it is hard to imagine someone uttering (138) without thereby addressing the question of whether the number rolled is prime:

(138)   If the number rolled is prime, it must be even.

For example, say that a fair die is rolled and placed under a red cup just in case it comes up 2, and placed under a green cup otherwise. Conversations about this case may address multiple questions, each corresponding to a different partition of logical space. (138) may address which cup the die is placed under. But in addition, (138) itself raises another question, namely whether a prime number was rolled. And as long as this question is in the algebra generated by the contextually salient partition, my theory predicts that (138) will sound bad to anyone with equal credence in each die outcome.[34] The same goes for many indicative conditionals, which commonly raise the question of whether their antecedents are acceptable.

To sum up: simple epistemic sentences are commonly used to address whether their prejacent is the case. This holds for 'probably' sentences, 'might' sentences, and indicative conditionals. Further investigation should explore whether similar generalizations hold for other epistemic vocabulary, such as strong modals and epistemic adjectives. In the end, further research

---

33 I am grateful to an anonymous referee for suggesting that my theory could deliver this prediction.

34 The unassertability of (138) follows from the fact that the resulting partition will contain elements that accept that the number rolled is prime, but not that it must be even.

should also decide whether such generalizations follow from semantic stipulations or from general pragmatic reasoning.

The choice between semantic and pragmatic explanations arises for other constructions involving epistemic vocabulary as well. I will conclude this paper by explaining some facts about strong modals that have been recently argued to be pragmatic in nature. Suppose that we are going to a party and wondering whether our friend Ted is already there. Consider the following minimal pair:

(139)   Ted is there.

(140)   Ted must be there.

It is not appropriate to utter or even merely suppose either of these sentences along with the negation of (139):

(141)   #Suppose that Ted is not there and that he is there.

(142)   #Suppose that Ted is not there and that he must be there.

This suggests that both (139) and (140) are semantically inconsistent with the negation of (139). In spite of this similarity, though, (140) feels less forceful than (139). For instance, 'your keys must be in the drawer' inspires less confidence than 'your keys are in the drawer' when we are looking for my keys. Furthermore, 'must' carries an evidential signal, namely that you have come to believe the prejacent as a result of some indirect inference. For example, (140) is inappropriate when you are staring at a party and see that Ted is present, but appropriate if you believe that Ted is at the party on the basis of seeing his car parked outside.[35]

These similarities and differences between (139) and (140) are explained by my theory of epistemic vocabulary. Recall that 'John is probably guilty' could address the simple question of whether John is guilty, or a more complicated question about how likely it is that John is guilty. In the same spirit, (139) and (140) could address the simple question of whether Ted is at the party, or a more complicated question about how likely it is that Ted is at the party. The relevant observation for our purposes is that compared with simple sentences, sentences containing epistemic vocabulary are more likely to address more complicated questions about likelihoods. The statement that

---

35 For further discussion of these observations, see von Fintel & Gillies 2010 and the earlier discussions they reference, including Karttunen 1972 and Groenendijk & Stokhof 1975.

Ted is at the party addresses the question of whether Ted is at the party, while the statement that Ted *must* be at the party intuitively addresses some question about the likelihood that Ted is at the party, affirming that the likelihood is indeed very high. This intuitive distinction neatly corresponds to a theoretical distinction: (139) prompts context to supply a coarse-grained partition to the evaluation of epistemic vocabulary, where the elements of the partition simply say whether Ted is at the party. (140) prompts context to supply a more fine-grained partition, where the elements of the partition correspond to evidence about whether Ted is at the party.

According to my semantics, (139) and (140) have the same effect on your credences, namely that you only have credence in partition elements that accept that Ted is at the party. But since (140) prompts the introduction of a more fine-grained partition, (140) demands that you are certain of the proposition that Ted is at the party *in virtue of being certain of some union of evidence propositions*, each of which definitely confirms that Ted is at the party. Hence my theory explains why (139) and (140) are both semantically inconsistent with the claim that Ted is not at the party, namely because both sentences demand that you give full credence to the proposition that Ted is at the party. At the same time, though, the assertion of (140) is the result of one or more inferences. The evidential component of 'must' corresponds to the confirmation of its prejacent by one or more elements of the contextually supplied partition. Following von Fintel & Gillies 2010, my theory avoids misinterpreting this evidential feature as genuine semantic weakness. Rather, 'must' sentences inspire less confidence than simple sentences because they highlight potential sources of doubt in their semantic values. The constraint expressed by a 'must' sentence may be rejected because you doubt that the relevant evidence propositions support the constraint, or because you doubt the evidence propositions themselves.

There are several other subjects that deserve investigation beyond the scope of this paper. For instance, we should explain why your credences may be contained in the content of an indicative conditional, even if you have no credence in its antecedent. That explanation may call for a more fine-grained model of your mental life, according to which you have primitive conditional credences. In addition, we have set aside cases where context contributes infinite partitions to the contents of sentences; such cases will ultimately demand modifications of the theory developed here. Furthermore, arguments in this paper do not settle exactly what it is to have a credence for practical purposes, or exactly what sorts of attitudes are constrained by assertions.

Further development of the theory should also discuss shared conversational features such as the common ground, saying whether these features have some probabilistic structure and saying how they are affected by probabilistic assertions (*cf.* Yalcin 2012b for commentary and progress on this project). And finally, a more complete theory should give even stronger guidelines for how context determines which partitions are used for interpreting epistemic expressions, perhaps even connecting these guidelines with independently motivated literature on questions under discussion (*cf.* Roberts 2012).

This paper is a progress report. The main goal of this paper is not to prove that the theory developed here is correct, nor that it is the only way to account for the observations we started with. The goals are more modest, namely to characterize the behavior of epistemic vocabulary and to develop a theory that makes sense of that behavior. The theoretical moral: epistemic vocabulary may concern not just your opinions about particular propositions, but more structured properties of your opinions. The empirical moral: context plays a role in determining exactly what structure is relevant. The semantic and pragmatic theories informed by these morals provide a unified account of several distinctive features of epistemic vocabulary. In addition to the specific theories defended here, I hope that the more general theoretical and empirical morals informing these theories may prove to be useful springboards for further research.

A   Lexical entries

For ease of reference, here is the semantics for the fragment I have discussed:

$$[\![\textbf{or}_i]\!]^c = [\lambda S.\lambda T.\{m: \forall p \in g_c(i), \ m|_p \in S \text{ or } m|_p \in T\}]$$

$$[\![\textbf{and}_i]\!]^c = [\lambda S.\lambda T.\{m: \forall p \in g_c(i), \ m|_p \in S \text{ and } m|_p \in T\}]$$

$$[\![\textbf{not}_i]\!]^c = [\lambda S. \quad \{m: \forall p \in g_c(i), \ m|_p \notin S\}]$$

$$[\![\textbf{might}_i]\!]^c = [\lambda S. \quad \{m: \exists p \in g_c(i) \text{ such that } m|_p \in S\}]$$

$$[\![\textbf{must}_i]\!]^c = [\lambda S. \quad \{m: \forall p \in g_c(i), \ m|_p \in S\}]$$

$$[\![\textbf{probably}_i]\!]^c = [\lambda S. \quad \{m: m\left(\bigcup\{p \in g_c(i): m|_p \in S\}\right) > {}^1\!/_2\}]$$

$$[\![\textbf{if}_i]\!]^c = [\lambda S.\lambda T.\{m: m\left(\bigcup\{p \in g_c(i): m|_p \in T\}\right.$$

$$\left. \Big| \bigcup\{p \in g_c(i): m|_p \in S\}\right) = 1\}]$$

$$[\![\mathcal{C}]\!]^c = [\lambda p.\{m: m(p) = 1\}]$$

In addition, as discussed Section 3, epistemic expressions and lower-type logical operators retain their traditional semantic values.

B   Derivations

In this section, I use 'high' as shorthand for 'the die landed on an high number', 'one' for 'the die landed on a one', and so on. Let us expand the natural language fragment under discussion in this paper with lexical entries such as $[\![\textbf{high}]\!] = \{w: \text{high in } w\}$, $[\![\textbf{one}]\!] = \{w: \text{one in } w\}$, and so on. Let c_0 be a context that resolves the values of referential variables as follows:

$$g_{c_0}(1) = \{[\![\textbf{low}]\!], [\![\textbf{high}]\!]\}$$

$$g_{c_0}(2) = \{[\![\textbf{one}]\!], [\![\textbf{two}]\!], [\![\textbf{three}]\!], [\![\textbf{four}]\!], [\![\textbf{five}]\!], [\![\textbf{six}]\!]\}$$

$$g_{c_0}(3) = \{\top\}$$

In Sections B.1–B.3, all semantic values are computed relative to c_0. Superscripts are omitted for readability. Let m_0 be the credence distribution you should have regarding how a certain die landed, if you are certain it is fair. In other words: $m_0([\![\textbf{one}]\!]) = {}^1\!/_6$, $m_0([\![\textbf{two}]\!]) = {}^1\!/_6$, and so on, and conditional on each proposition about how the die landed, m_0 assigns probability 1 to the correct proposition about its parity and size, counting numbers up to three as low and numbers above three as high. This credence distribution will function as our countermodel for the invalidity proofs in Sections B.2–B.3.

B.1 An example of 'probably' under disjunction

We demonstrate that m_0 is contained in the semantic value of (40):

(40)   [probably$_2$ $\mathcal{C}$ it is even] or$_1$ [$\mathcal{C}$ less than four]

The precise semantic value of (40) is as follows:

$$\llbracket(40)\rrbracket = \{m\colon \forall p \in g_{c_0}(1),$$
$$m|_p \in \llbracket\mathbf{probably_2}\ \mathcal{C}\ \mathbf{it\ is\ even}\rrbracket\ \text{or}$$
$$m|_p \in \llbracket\mathcal{C}\ \mathbf{it\ is\ less\ than\ four}\rrbracket\}$$

$$= \{m\colon \forall p \in g_{c_0}(1),$$
$$m|_p \in \{m'\colon m'\big(\bigcup\{p' \in g_{c_0}(2)\colon m'|_{p'} \in \llbracket\mathcal{C}\ \mathbf{it\ is\ even}\rrbracket\}\big) > \text{\textonehalf}\}\ \text{or}$$
$$m|_p \in \llbracket\mathcal{C}\ \mathbf{it\ is\ less\ than\ four}\rrbracket\}$$

$$= \{m\colon \forall p \in g_{c_0}(1),$$
$$m|_{p'} \in \{m'\colon m'\big(\bigcup\{p' \in g_{c_0}(2)\colon$$
$$m'|_p \in \{m''\colon m''(\{w\colon \text{it is even in } w\}) = 1\}\}\big) > \text{\textonehalf}\}\ \text{or}$$
$$m|_p \in \{m'''\colon m'''(\{w\colon \text{it is less than four in } w\}) = 1\}\}$$

The following facts follow from the construction of m_0:

$$m_0|_{\llbracket\mathbf{low}\rrbracket}(\llbracket\mathbf{less\ than\ four}\rrbracket) = 1$$
$$m_0|_{\llbracket\mathbf{high}\rrbracket}|_{\llbracket\mathbf{four}\rrbracket}(\llbracket\mathbf{even}\rrbracket) = 1$$
$$m_0|_{\llbracket\mathbf{high}\rrbracket}|_{\llbracket\mathbf{six}\rrbracket}(\llbracket\mathbf{even}\rrbracket) = 1$$

And so we may conclude:

$$m_0|_{\llbracket\mathbf{low}\rrbracket} \in \llbracket\mathcal{C}\ \mathbf{it\ is\ less\ than\ four}\rrbracket$$
$$m_0|_{\llbracket\mathbf{high}\rrbracket}|_{\llbracket\mathbf{four}\rrbracket} \in \llbracket\mathcal{C}\ \mathbf{it\ is\ even}\rrbracket$$
$$m_0|_{\llbracket\mathbf{high}\rrbracket}|_{\llbracket\mathbf{six}\rrbracket} \in \llbracket\mathcal{C}\ \mathbf{it\ is\ even}\rrbracket$$

Furthermore, $m_0|_{\llbracket\mathbf{high}\rrbracket}(\llbracket\mathbf{four}\rrbracket \bigcup \llbracket\mathbf{six}\rrbracket) = \text{\textexclamdown}2/3 > \text{\textonehalf}$. As a result, since $g_{c_0}(2)$ contains both $\llbracket\mathbf{four}\rrbracket$ and $\llbracket\mathbf{six}\rrbracket$, we can conclude that $m_0|_{\llbracket\mathbf{high}\rrbracket} \in \llbracket\mathbf{probably_2}\ \mathcal{C}\ \mathbf{it\ is\ even}\rrbracket$. Finally, since $g_{c_0}(1)$ contains just $\llbracket\mathbf{low}\rrbracket$ and $\llbracket\mathbf{high}\rrbracket$, it follows that m_0 is contained in the content of the disjunction (40) itself.

B.2 The invalidity of modus tollens

We demonstrate that the following argument is invalid:

(143) a. if$_1$ $\mathcal{C}$ low, probably$_2$ $\mathcal{C}$ odd
      b. not$_3$ probably$_2$ $\mathcal{C}$ odd
      c. Hence: $\mathcal{C}$ not low

In what follows, we show that $m_0 \in [\![(143a)]\!]$, $m_0 \in [\![(143b)]\!]$, and $m_0 \notin [\![(143c)]\!]$. First, note that $[\![(143a)]\!] = \{m: m(\mathcal{B}|\mathcal{A}) = 1\}$, where $\mathcal{A}$ and $\mathcal{B}$ are defined as follows:

$$\mathcal{A} = \bigcup \{p \in g_{c_0}(1): m|_p \in [\![\mathcal{C}\ \mathbf{low}]\!]\}$$
$$\mathcal{B} = \bigcup \{p \in g_{c_0}(1): m|_p \in [\![\mathbf{probably}_2\ \mathcal{C}\ \mathbf{odd}]\!]\}$$

First we wish to test whether $m_0 \in [\![(143a)]\!]$. The answer depends on $m_0(\mathcal{B}|\mathcal{A})$. The proposition $\mathcal{A}$ is the following union of propositions:

$$\mathcal{A} = \bigcup \{p \in g_{c_0}(1): m_0|_p \in [\![\mathcal{C}\ \mathbf{low}]\!]\}$$
$$= \bigcup \{p \in \{[\![\mathbf{low}]\!], [\![\mathbf{high}]\!]\}: m_0|_p \in [\![\mathcal{C}\ \mathbf{low}]\!]\}$$
$$= \bigcup \{p \in \{[\![\mathbf{low}]\!], [\![\mathbf{high}]\!]\}: m_0|_p \in \{m': m'([\![\mathbf{low}]\!]) = 1\}\}$$

Since $(m_0|_{[\![\mathbf{low}]\!]})([\![\mathbf{low}]\!]) = 1$, $[\![\mathbf{low}]\!]$ is a member of the set of propositions whose union is $\mathcal{A}$. But since $(m_0|_{[\![\mathbf{high}]\!]})([\![\mathbf{low}]\!]) \neq 1$, $[\![\mathbf{high}]\!]$ is not a member of that set. Since $[\![\mathbf{low}]\!]$ and $[\![\mathbf{high}]\!]$ are the only candidate members of the set whose union is $\mathcal{A}$, we may conclude that $\mathcal{A} = [\![\mathbf{low}]\!]$. Similarly, $\mathcal{B}$ is the following union of propositions:

$$\mathcal{B} = \bigcup \{p \in g_{c_0}(1): m_0|_p \in [\![\mathbf{probably}_2\ \mathcal{C}\ \mathbf{odd}]\!]\}, \text{ and}$$
$$= \bigcup \{p \in \{[\![\mathbf{low}]\!], [\![\mathbf{high}]\!]\}: m_0|_p \in [\![\mathbf{probably}_2\ \mathcal{C}\ \mathbf{odd}]\!]\}$$
$$= \bigcup \{p \in \{[\![\mathbf{low}]\!], [\![\mathbf{high}]\!]\}:$$
$$m_0|_p \in \{m': m'\Big(\bigcup \{p' \in \{[\![\mathbf{one}]\!], [\![\mathbf{two}]\!], \ldots, [\![\mathbf{six}]\!]\}:$$
$$m'|_{p'} \in \{m'': m''([\![\mathbf{odd}]\!]) = 1\}\}\Big) > {}^1\!/{}_2\}\}$$

Since $p' = [\![\mathbf{one}]\!]$ and $p' = [\![\mathbf{three}]\!]$ are among the values for which $(m_0|_{[\![\mathbf{low}]\!]})|_{p'}([\![\mathbf{odd}]\!]) = 1$, and we have $m_0|_{[\![\mathbf{low}]\!]}([\![\mathbf{one}]\!] \bigcup [\![\mathbf{three}]\!]) = {}^2\!/{}_3 > {}^1\!/{}_2$, we may conclude that $[\![\mathbf{low}]\!]$ is in the set of propositions whose union is $\mathcal{B}$. Furthermore, since the values for which $(m_0|_{[\![\mathbf{high}]\!]})|_{p'}([\![\mathbf{odd}]\!]) = 1$ include at most the values $p' = [\![\mathbf{one}]\!], [\![\mathbf{three}]\!], [\![\mathbf{five}]\!]$, and we have $m_0|_{[\![\mathbf{high}]\!]}([\![\mathbf{one}]\!] \bigcup [\![\mathbf{three}]\!] \bigcup [\![\mathbf{five}]\!]) = {}^1\!/{}_3 \not> {}^1\!/{}_2$, we may conclude that $[\![\mathbf{high}]\!]$ is not in the set of propositions whose union is $\mathcal{B}$. Since $[\![\mathbf{low}]\!]$ and $[\![\mathbf{high}]\!]$

are the only candidate members of the set whose union is $\mathcal{B}$, we may conclude that $\mathcal{B} = [\![\textbf{low}]\!]$. Finally, since $m_0([\![\textbf{low}]\!] | [\![\textbf{low}]\!]) = 1$, it follows that $m_0 \in [\![(143\text{a})]\!]$.

Second, we compute the semantic value $[\![(143\text{b})]\!]$:

$$
\begin{aligned}
[\![(143\text{b})]\!] &= \{m \colon \forall p \in g_{c_0}(3),\ m|_p \notin [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!]\} \\
&= \{m \colon \forall p \in \{\top\},\ m|_p \notin [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!]\} \\
&= \{m \colon m \notin [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!]\} \\
&= \{m \colon m \notin \{m' \colon m'\big(\bigcup\{p \in g_{c_0}(2) \colon m'|_p \in [\![\mathcal{C}\ \textbf{odd}]\!]\}\big) > {}^1\!/_2\}\} \\
&= \{m \colon m \notin \{m' \colon m'\big(\bigcup\{p \in g_{c_0}(2) \colon m'|_p([\![\textbf{odd}]\!]) = 1\}\big) > {}^1\!/_2\}\} \\
&= \{m \colon m \notin \{m' \colon m'\big(\bigcup\{p \in \{[\![\textbf{one}]\!], [\![\textbf{two}]\!], \ldots, [\![\textbf{six}]\!]\} \colon \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad m'|_p([\![\textbf{odd}]\!]) = 1\}\big) > {}^1\!/_2\}\}
\end{aligned}
$$

Since $p = [\![\textbf{one}]\!], [\![\textbf{three}]\!], [\![\textbf{five}]\!]$ are the values for which $m_0|_p([\![\textbf{odd}]\!]) = 1$, and we have $m_0([\![\textbf{one}]\!] \cup [\![\textbf{three}]\!] \cup [\![\textbf{five}]\!]) = {}^1\!/_2 \not> {}^1\!/_2$, we have that $m_0 \notin [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!]$, and it follows that $m_0 \in [\![(143\text{b})]\!]$.

Third, we compute $[\![(143\text{c})]\!]$:

$$
\begin{aligned}
[\![(143\text{c})]\!] &= \{m \colon m([\![\textbf{not low}]\!]) = 1\} \\
&= \{m \colon m(\{w \colon w \notin [\![\textbf{low}]\!]\}) = 1\}
\end{aligned}
$$

But $m_0(\{w \colon w \notin [\![\textbf{low}]\!]\}) = {}^1\!/_2 \neq 1$, so $m_0 \notin [\![(143\text{c})]\!]$. Hence the argument from $[\![(143\text{a})]\!]$ and $[\![(143\text{b})]\!]$ to $[\![(143\text{c})]\!]$ is not valid in the context c_0, and therefore it is not valid simpliciter.

B.3 The invalidity of constructive dilemma

We demonstrate that the following argument is invalid:

(144) a. if$_1$ $\mathcal{C}$ low, probably$_2$ $\mathcal{C}$ odd
  b. if$_1$ $\mathcal{C}$ high, probably$_2$ $\mathcal{C}$ even
  c. $\mathcal{C}$ [low or high]
  d. Hence: [probably$_2$ $\mathcal{C}$ odd] or$_3$ [probably$_2$ $\mathcal{C}$ even]

As demonstrated in Section B.2, we have $m_0 \in [\![(144\text{a})]\!]$. An analogous argument demonstrates that $m_0 \in [\![(144\text{b})]\!]$, if we simply replace 'low', 'high', 'odd', 'even', 'one', 'two', 'three', 'four', 'five', and 'six' as they occur in the

argument with 'high', 'low', 'even', 'odd', 'four', 'five', 'six', 'one', 'two', and 'three', respectively. It just remains to be shown that $m_0 \in [\![(144\text{c})]\!]$ and $m_0 \notin [\![(144\text{d})]\!]$.

First, we compute the semantic value $[\![(144\text{c})]\!]$:

$$
\begin{aligned}
[\![(144\text{c})]\!] &= \{m \colon m([\![\textbf{low or high}]\!]) = 1\} \\
&= \{m \colon m(\{w \colon w \in [\![\textbf{low or high}]\!]\}) = 1\} \\
&= \{m \colon m(\{w \colon w \in [\![\textbf{low}]\!] \text{ or } w \in [\![\textbf{high}]\!]\}) = 1\}
\end{aligned}
$$

Since $m_0(\{w \colon w \in [\![\textbf{low}]\!] \text{ or } w \in [\![\textbf{high}]\!]\}) = 1$, it follows that $m_0 \in [\![(144\text{c})]\!]$.

Second, we compute the semantic value $[\![(144\text{d})]\!]$:

$$
\begin{aligned}
[\![(144\text{d})]\!] &= \{m \colon \forall p \in g_{c_0}(3),\ m|_p \in [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!] \text{ or} \\
&\qquad\qquad\qquad m|_p \in [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{even}]\!]\} \\
&= \{m \colon \forall p \in \{\top\},\ m|_p \in [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!] \text{ or} \\
&\qquad\qquad\qquad m|_p \in [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{even}]\!]\} \\
&= \{m \colon m \in [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!] \text{ or} \\
&\qquad\qquad m \in [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{even}]\!]\}
\end{aligned}
$$

But recall from Section B.2 that $m_0 \notin [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{odd}]\!]$, and an analogous argument demonstrates that $m_0 \notin [\![\textbf{probably}_2\ \mathcal{C}\ \textbf{even}]\!]$. From this it follows that $m_0 \notin [\![(144\text{d})]\!]$.

B.4   The quasi-validity of modus tollens

Formally, an argument A is *quasi-valid* just in case: for any measure m and any context c such that for every index i on epistemic vocabulary in A we have $m(p) > 0$ for all $p \in g_c(i)$, if m is contained in the semantic values of the premises of A in c, then m is contained in the semantic value of the conclusion of A in c.

We can demonstrate that the following argument schema is quasi-valid, *i.e.* that replacing each schematic letter with a sentence whose semantic value is a set of measures always yields a quasi-valid argument:

(145)   a.   if$_1$ P, Q
       b.   not$_1$ Q
       c.   Hence: not$_1$ P

Let m be an arbitrary measure contained in both $\llbracket(145a)\rrbracket$ and $\llbracket(145b)\rrbracket$, and let c be an arbitrary context such that $m(p) > 0$ for all $p \in g_c(1)$. Since $m \in \llbracket(145a)\rrbracket$, we know that $m(\mathcal{Q}|\mathcal{P}) = 1$, where we have:

$$\mathcal{P} = \bigcup\{p \in g_c(1): m|_p \in \llbracket\mathbf{P}\rrbracket\}$$
$$\mathcal{Q} = \bigcup\{p \in g_c(1): m|_p \in \llbracket\mathbf{Q}\rrbracket\}$$

Furthermore, since $m \in \llbracket(145b)\rrbracket$, the set $\{p \in g_c(1): m|_p \in \llbracket\mathbf{Q}\rrbracket\}$ is empty. Hence $\mathcal{Q} = \bot$. For reductio, assume that $m \notin \llbracket(145c)\rrbracket$. In other words, $m \notin \{m: \forall p \in g_c(1), m|_p \notin \llbracket\mathbf{P}\rrbracket\}$. From this it follows that the set $\{p \in g_c(1): m|_p \in \llbracket\mathbf{P}\rrbracket\}$ is not empty, hence $\mathcal{P} \neq \bot$. Since $m(p) > 0$ for all $p \in g_c(1)$, we can infer that $m(\mathcal{Q}|\mathcal{P})$ is well-defined, and since $\mathcal{Q} = \bot$, we can infer that $m(\mathcal{Q}|\mathcal{P}) = 0 \neq 1$. But this contradicts our assumption that $m \in \llbracket(145a)\rrbracket$.

B.5 The quasi-validity of constructive dilemma

We can demonstrate that the following argument schema is quasi-valid:

(146) a. if$_1$ P, Q
       b. if$_1$ R, S
       c. P or$_1$ R
       d. Hence: Q or$_1$ S

Let m be an arbitrary measure contained in $\llbracket(146a)\rrbracket$, $\llbracket(146b)\rrbracket$, and $\llbracket(146c)\rrbracket$, and let c be an arbitrary context such that $m(p) > 0$ for all $p \in g_c(1)$. Since $m \in \llbracket(146a)\rrbracket$, we know that $m(\mathcal{Q}|\mathcal{P}) = 1$, where $\mathcal{P}$ and $\mathcal{Q}$ are defined as in B.4. Hence $m(\mathcal{P}) = m(\mathcal{PQ})$. Furthermore, since m is a probability measure and $g_c(1)$ is a partition, we have:

$$m(\mathcal{P}) = \sum_{\substack{p \in g_c(1) \\ p \subseteq \mathcal{P}}} m(p) = \sum_{\substack{p \in g_c(1) \\ p \subseteq \mathcal{Q} \\ p \subseteq \mathcal{P}}} m(p) + \sum_{\substack{p \in g_c(1) \\ p \nsubseteq \mathcal{Q} \\ p \subseteq \mathcal{P}}} m(p) = m(\mathcal{PQ}) + \sum_{\substack{p \in g_c(1) \\ p \nsubseteq \mathcal{Q} \\ p \subseteq \mathcal{P}}} m(p)$$

From this we may conclude that there is no $p \in g_c(1)$ such that $p \nsubseteq \mathcal{Q}$, $p \subseteq \mathcal{P}$, and $m(p) > 0$. Since $m(p) > 0$ for all $p \in g_c(1)$ by stipulation, it follows that for all $p \in g_c(1)$, if $p \subseteq \mathcal{P}$ then $p \subseteq \mathcal{Q}$. By an analogous argument, the fact that $m \in \llbracket(146b)\rrbracket$ entails that for all $p \in g_c(1)$, if $p \subseteq \mathcal{R}$ then $p \subseteq \mathcal{S}$,

where we have:

$$\mathcal{R} = \bigcup \{p \in g_c(1) : m|_p \in [\![\mathbf{R}]\!]\}$$
$$\mathcal{S} = \bigcup \{p \in g_c(1) : m|_p \in [\![\mathbf{S}]\!]\}$$

Since $m \in [\![(146c)]\!]$, for all $p \in g_c(1)$, either $p \subseteq \mathcal{P}$ or $p \subseteq \mathcal{R}$. This taken together with the results just proved entails that for all $p \in g_c(1)$, either $p \subseteq \mathcal{Q}$ or $p \subseteq \mathcal{S}$, from which it follows that $m \in [\![(146d)]\!]$, as desired.

B.6 The validity of chancy modus ponens

We can demonstrate that the following argument schema is valid:

(147) a. probably$_1$ P
     b. if$_1$ P, Q
     c. Hence: probably$_1$ Q

Let m be an arbitrary measure contained in $[\![(147a)]\!]$ and $[\![(147b)]\!]$, and let us define $\mathcal{P}$ and $\mathcal{Q}$ as in Appendix B.4. Since $m \in [\![(147a)]\!]$, we know that $m(\mathcal{P}) > 1/2$. Furthermore, since $m(\mathcal{P}) > 0$, and since $m \in [\![(147b)]\!]$ entails that $m(\mathcal{Q}|\mathcal{P}) = 1$, we can infer that $m(\mathcal{PQ}) = m(\mathcal{P})$. And finally, since m is a probability measure, we know that $m(\mathcal{Q}) \geq m(\mathcal{PQ})$. To sum up: we have that $m(\mathcal{Q}) \geq m(\mathcal{PQ}) = m(\mathcal{P}) > 1/2$, from which it follows that $m \in [\![(147c)]\!]$.

References

Austen, Jane. 1818. *Persuasion*. Ed. Patricia Meyer Spacks. Norton Critical Editions. 1st edition, printed in 1995. New York: Norton.

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

Bennett, Jonathan. 2003. *A philosophical guide to conditionals*. Oxford: Oxford University Press. http://dx.doi.org/10.1093/0199258872.001.0001.

Cantwell, John. 2008. Changing the modal context. *Theoria* 74(4). 331–51. http://dx.doi.org/10.1111/j.1755-2567.2008.00028.x.

Cappelen, Herman & John Hawthorne. 2009. *Relativism and monadic truth*. Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199560554.001.0001.

Carroll, Lewis. 1894. A logical paradox. *Mind* 3(11). 436–40.

Condoravdi, Cleo & Jean Gawron. 1996. The context-dependency of implicit arguments. In Makoto Kanazawa, Christopher Piñón & Henriëtte de Swart (eds.), *Quantifiers, deduction, and context*, 1–32. Stanford: CSLI Publications.

Dever, Josh. 2012. Must or might. Ms., University of Texas at Austin: https://webspace.utexas.edu/deverj/personal/papers/mustAPA.pdf.

Diaconis, Persi & Sandy L. Zabell. 1982. Updating subjective probability. *Journal of the American Statistical Association* 77. 822–830.

Dorr, Cian & John Hawthorne. 2012. Embedding epistemic modals. Ms., Department of Philosophy, Oxford University.

Edgington, Dorothy. 1995. On conditionals. *Mind* 104. 235–329. http://dx.doi.org/10.1093/mind/104.414.235.

Egan, Andy. 2007. Epistemic modals, relativism, and assertion. *Philosophical Studies* 133(1). 1–22. http://dx.doi.org/10.1007/s11098-006-9003-x.

Egan, Andy, John Hawthorne & Brian Weatherson. 2005. Epistemic modals in context. In Gerhard Preyer & Georg Peter (eds.), *Contextualism in philosophy*. Oxford: Oxford University Press.

Egan, Andy & Brian Weatherson (eds.). 2011. *Epistemic modality*. Oxford: Oxford University Press.

Fantl, Jeremy & Matthew McGrath. 2010. *Knowledge in an uncertain world*. Oxford: Oxford University Press. http://dx.doi.org/11.1093/acprof:oso/9780199550623.001.0001.

von Fintel, Kai & Anthony S. Gillies. 2008a. An opinionated guide to epistemic modality. In Tamar Szabó Gendler & John Hawthorne (eds.), *Oxford studies in epistemology*, vol. 2, 32–62. Oxford: Oxford University Press. http://dx.doi.org/10.1215/00318108-2007-025.

von Fintel, Kai & Anthony S. Gillies. 2008b. CIA leaks. *Philosophical Review* 117(1). 77–98. http://dx.doi.org/10.1215/00318108-2007-025.

von Fintel, Kai & Anthony S. Gillies. 2010. 'Must'... stay...strong! *Natural Language Semantics* 18(4). 351–83. http://dx.doi.org/10.1007/s11050-010-9058-2.

von Fintel, Kai & Anthony S. Gillies. 2011. *Might* made right. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 108–30. Oxford: Oxford University Press.

Ganson, Dorit. 2008. Evidentialism and pragmatic constraints on outright belief. *Philosophical Studies* 139(3). 441–58. http://dx.doi.org/10.1007/s11098-007-9133-9.

Gazdar, Gerald. 1980. A cross-categorial semantics for coordination. *Linguistics and Philosophy* 3(3). 407–9. http://dx.doi.org/10.1007/BF00401693.

Geurts, Bart. 1999. *Presuppositions and pronouns*. Amsterdam: Elsevier.

Gibbard, Allan. 1981. Two recent theories of conditionals: Conditionals, belief, decision, chance, and time. In William L. Harper, Robert Stalnaker & Glenn Pearce (eds.), *Ifs: Conditionals, belief, decision, chance, and time*, 211–47. Dordrecht: D. Reidel Publishing Company.

Gillies, Anthony S. 2004. Epistemic conditionals and conditional epistemics. *Noûs* 38(4). 585–616. http://dx.doi.org/10.1111/j.0029-4624.2004.00485.x.

Gillies, Anthony S. 2010. Iffiness. *Semantics and Pragmatics* 3(4). 1–42. http://dx.doi.org/10.3765/sp.3.4.

Groenendijk, Jeroen & Martin Stokhof. 1975. Modality and conversational information. *Theoretical Linguistics* 2. 61–112. http://dx.doi.org/10.1515/thli.1975.2.1-3.61.

Hacquard, Valentine & Alexis Wellwood. 2012. Embedding epistemic modals in English: A corpus-based study. *Semantics and Pragmatics* 5(4). 1–29. http://dx.doi.org/10.3765/sp.5.4.

Hart, H. L. A. & Tony Honoré. 1985. *Causation in the law*. 2nd ed. Oxford: Oxford University Press.

Hawthorne, John. 2007. Eavesdroppers and epistemic modals. *Philosophical Issues* 17(1). 92–101. http://dx.doi.org/10.1111/j.1533-6077.2007.00124.x.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. http://dx.doi.org/10.1093/jos/9.3.183.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell Publishers, Ltd.

Hendriks, Petra. 2004. *Either*, *both* and *neither* in coordinate structures: From lexeme to discourse. In Alice ter Meulen & Werner Abraham (eds.), *The composition of meaning: From lexeme to discourse*, 115–138. Amsterdam: John Benjamins.

Joyce, James. 2005. How probabilities reflect evidence. *Philosophical Perspectives* 19. 153–78.

Karttunen, Lauri. 1972. *Possible* and *must*. In J. Kimball (ed.), *Syntax and semantics*, vol. 1, 1–20. New York: Academic Press.

Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4(2). 169–93.

Kaufmann, Stefan. 2004. Conditioning against the grain. *Journal of Philosophical Logic* 33(6). 583–606. http://dx.doi.org/10.1023/B:LOGI.0000046142.51136.bf.

Knobe, Joshua & Seth Yalcin. 2015. Fat Tony might be dead: An experimental note on epistemic modals. *Semantics and Pragmatics* 7(10). 1–21. http://dx.doi.org/10.3765/sp.7.10.

Kolodny, Niko & John MacFarlane. 2010. Ifs and oughts. *Journal of Philosophy* 107(3). 115–43.

Kratzer, Angelika. 1977. What *must* and *can* must and can mean. *Linguistics and Philosophy* 1(3). 337–55. http://dx.doi.org/10.1007/BF00353453.

Kratzer, Angelika. 1981. The notional category of modality: New approaches in word semantics. In Hans Jurgen Eikmeyer & Hannes Rieser (eds.), *Words, worlds, and contexts: New approaches in word semantics*, 38–74. Berlin: W. de Gruyter.

Kratzer, Angelika. 1991. Modality: An international handbook of contemporary research. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–50. Berlin: W. de Gruyter.

Landman, Fred. 1986. Conflicting presuppositions and modal subordination. *Chicago Linguistics Society* 22. 195–207.

Lasersohn, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28(6).

Lycan, William. 2001. *Real conditionals*. Oxford: Oxford University Press. http://dx.doi.org/10.1080/713659573.

MacFarlane, John. 2011. Epistemic modals are assessment-sensitive. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 144–78. Oxford: Oxford University Press.

Mackie, J. L. 1974. *The cement of the universe: A study of causation*. Oxford: Oxford University Press.

McGee, Vann. 1985. A counterexample to modus ponens. *Journal of Philosophy* 82(9). 462–71. http://dx.doi.org/10.2307/2026276.

Moss, Sarah. 2013. Epistemology formalized. *Philosophical Review* 122(1). 1–43. http://dx.doi.org/10.1215/00318108-1728705.

Moss, Sarah. 2014. *Probability and knowledge*. Book manuscript. Department of Philosophy, University of Michigan.

Partee, Barbara. 1986. Noun phrase interpretation and type-shifting principles. In *Compositionality in formal semantics*, 203–30. Oxford: Blackwell Publishers, Ltd. http://dx.doi.org/10.1002/9780470751305.

Partee, Barbara. 1989. Binding implicit variables in quantified contexts. In *Compositionality in formal semantics*, 259–81. Oxford: Blackwell Publishers, Ltd. http://dx.doi.org/10.1002/9780470751305.

Partee, Barbara. 2004. *Compositionality in formal semantics*. Oxford: Blackwell Publishers, Ltd. http://dx.doi.org/10.1002/9780470751305.

Partee, Barbara & Mats Rooth. 1982. Conjunction, type-ambiguity, and wide scope 'or'. *Proceedings of the First West Coast Conference on Formal Linguistics*. 1–10.

Partee, Barbara & Mats Rooth. 1983. Generalized conjunction and type-ambiguity. In Rainer Bauerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and the interpretation of language*, 361–83. Berlin: de Gruyter.

Roberts, Craige. 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12(6). 683–721. http://dx.doi.org/10.1007/BF00632602.

Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. http://dx.doi.org/10.3765/sp.5.6.

Rothschild, Daniel. 2011. Explaining presupposition projection with dynamic semantics. *Semantics and Pragmatics* 4(3). 1–43. http://dx.doi.org/10.3765/sp.4.3.

Rothschild, Daniel. 2012. Expressing credences. *Proceedings of the Aristotelian Society* 112(1). 99–114. http://dx.doi.org/10.1111/j.1467-9264.2012.00327.x.

Schaffer, Jonathan. 2005. Contrastive causation. *Philosophical Review* 114(3). 297–328. http://dx.doi.org/10.1215/00318108-114-3-327.

Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 2(3). 1–78. http://dx.doi.org/10.3765/sp.2.3.

Schroeder, Mark. 2012. Attitudes and epistemics. Ms., Department of Philosophy, University of Southern California.

Schroeder, Mark & Jake Ross. 2014. Belief, credence, and pragmatic encroachment. *Philosophy and Phenomenological Research* 88(2). 259–88. http://dx.doi.org/10.1111/j.1933-1592.2011.00552.x.

Slote, M. 1978. Time in counterfactuals. *Philosophical Review* 87. 3–27. http://dx.doi.org/10.2307/2184345.

Stalnaker, Robert. 1970. Pragmatics. In *Context and content*, 31–46. Oxford: Oxford University Press. http://dx.doi.org/10.1007/s11098-010-9587-z.

Stalnaker, Robert. 2010. Responses to Stanley and Schlenker. *Philosophical Studies* 151(1). 143–57.

Swanson, Eric. 2006. Interactions with context. PhD. dissertation, Department of Linguistics and Philosophy, MIT.

Swanson, Eric. 2012. The application of constraint semantics to the language of subjective uncertainty. Forthcoming in the *Journal of Philosophical Logic*.

Veltman, Frank. 1985. *Logics for conditionals*. University of Amsterdam PhD thesis.

Veltman, Frank. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25(3). 221–61. http://dx.doi.org/10.1007/BF00248150.

Weatherson, Brian. 2005. Can we do without pragmatic encroachment? *Philosophical Perspectives* 19(1). 417–43. http://dx.doi.org/10.1111/j.1520-8583.2005.00068.x.

Weatherson, Brian. 2008. Attitudes and relativism. *Philosophical Perspectives* 22(1). 527–44.

Willer, Malte. 2013. Dynamics of epistemic modality. *Philosophical Review* 122(1). 45–92.

Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford: Oxford University Press. http://dx.doi.org/10.1215/00318108-1728714.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. http://dx.doi.org/10.1093/mind/fzm983.

Yalcin, Seth. 2009. More on epistemic modals. *Mind* 118. 785–93. http://dx.doi.org/10.1093/mind/fzp106.

Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11). 916–937. http://dx.doi.org/10.1111/j.1747-9991.2010.00360.x.

Yalcin, Seth. 2011. Nonfactualism about epistemic modality. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 295–332. Oxford: Oxford University Press.

Yalcin, Seth. 2012a. A counterexample to modus tollens. *Journal of Philosophical Logic* 41. 1001–24. http://dx.doi.org/10.1007/s10992-012-9228-4.

Yalcin, Seth. 2012b. Bayesian expressivism. *Proceedings of the Aristotelian Society* 112(2). 123–60. http://dx.doi.org/10.1111/j.1467-9264.2012.00329.x.

Sarah Moss
University of Michigan
Department of Philosophy
2215 Angell Hall, 435 South State Street
Ann Arbor, MI 48109
ssmoss@umich.edu