REASONS AND MODALS

by

Adam Marushak

B.A. in Philosophy, Boston University, 2005

M.A. in Philosophy, University of Wisconsin-Milwaukee, 2008

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

PhD in Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Adam Marushak

It was defended on

March 16, 2018

and approved by

Anil Gupta

James R. Shaw

Michael Caie

Karl Schafer

Eric Swanson

Dissertation Advisors: Anil Gupta,

James R. Shaw

REASONS AND MODALS

Adam Marushak, PhD

University of Pittsburgh, 2018

Reasons have become a central topic in philosophy. Epistemologists study reasons for belief. Moral philosophers assess reasons for action. And numerous metaethicists argue that all normative notions can ultimately be analyzed in terms of the concept of a reason. However, philosophers are not the only ones who talk about reasons. The Corpus of Contemporary American English yields over 10,000 examples of sentences containing the phrase "reason to [verb]", e.g. "reason to believe" and "reason to do". What, then, is the meaning of reason claims in colloquial language? This dissertation offers a theory of meaning for colloquial talk of reasons, focusing on sentences of the form "There is reason to believe that P". I argue that claims about reasons are a type of modal language. Familiar modals like "ought", "might", and "must" describe how things stand with relevant bodies of information, e.g. "You must pay your taxes" describes the relevant laws as requiring you to pay your taxes. I show that reason claims describe relevant information in an analogous manner: "There is reason to believe that P" describes the relevant knowledge as counting in favor of believing that P. This theory has far-reaching implications for recent debates about reasons and modals. First, the language of reasons pressures us to revise widely held views about the meaning of epistemic modals like "might": I show that it is talk of reasons for belief that describes knowledge, not epistemic modal language. Second, I argue that this semantic fact is best explained by a conceptual thesis: our pre-theoretical concept of a reason for belief is that of an item of knowledge, not a mere fact. Finally, once we see that "epistemic" modals have no special connection to knowledge, we are in a position to diagnose much of the confusion surrounding the formulation of fallibilism in epistemology. I argue that it is a mistake to

rely on intuitions about epistemic modals to assess the truth of fallibilism, and I propose an alternative methodology for determining whether fallibilism is true.

TABLE OF CONTENTS

PREFACE

I wish to thank all of those who have aided and inspired me during my years of studying philosophy. Special thanks are due to my advisors, Anil Gupta and James R. Shaw, for their dedication and support at every turn. Michael Caie, Karl Schafer, and Eric Swanson provided excellent guidance throughout the project, and my fellow Pitt graduate students gave me invaluable feedback. I thank also the faculty and graduate students at the University of Wisconsin-Milwaukee for helping me make the jump to graduate studies. I was fortunate to have outstanding mentors during my undergraduate studies as well, including Juliet Floyd, Klaus Brinkmann, and especially Alfred Tauber. I would never have studied philosophy without the inspiration of several key individuals during my highschool years. Tracy Beck-Briggs and Jerry Cossel nurtured my early interest in philosophy, and Kyle Christoff first introduced me to the discipline. Finally, I owe everything to my parents, Bernadette and Donald Marushak, who always supported me even as I pursued an unusual path through life. I dedicate this dissertation to them.

1.0 INTRODUCTION

The past several decades have witnessed a surge of philosophical interest in reasons. Epistemologists investigate the nature of reasons for belief.[1] Ethicists investigate the nature of reasons for action.[2] And numerous metaethicists argue that all normative notions can ultimately be analyzed in terms of the concept of a reason.[3]

The jumping-off point of this dissertation is the observation that philosophers are not the only ones who talk about reasons: claims about reasons are commonplace in colloquial language. For example, the Corpus of Contemporary American English yields over 10,000 examples of sentences containing the phrase "reason to [verb]", including phrases like "reason to believe", "reason to do", and "reason to feel". Languages ranging from Cantonese, to German, to Spanish all contain roughly synonymous expressions.

The prevalence of reason-talk in colloquial language raises the following question: How should we theorize about the meaning of reason claims in ordinary language? Despite the recent interest in reasons, little work has been done to address this question. Philosophers have devoted a great deal of attention to the meaning of normative terms like "ought", but the parallel investigation into the semantics of reasons is still at an early stage.

Let me be clear about the distinctively *semantic* investigation I undertake in this dissertation. The recent literature on reasons is rife with proposals for what we might refer to as the *metaphysics* of reasons. The questions at issue here are:

- What is the ontology of reasons? Are reasons facts, mental states, propositions, or something else?

[1] See, for example, the papers collected in Reisner and Steglich-Petersen (2011).
[2] See, for example, the papers collected in Sobel and Wall (2009).
[3] See, for example, Parfit (2011), Scanlon (1998, 2014), and Skorupski (2010).

- What makes something a reason for an act or an attitude? Is it possible to give a reductive analysis of the reason-for relation? And just how many relata are there in the reason-for relation?
- Is the concept of a reason the most fundamental normative concept? Or should we understand reasons in terms of notions like obligation, rationality, or fittingness?

Contrast the questions above with those concerning the *semantics* of reasons:

- Where should we locate natural language talk of reasons in a broad linguistic taxonomy? Are claims about reasons simply ordinary predications, or do they have some other type of structure?
- What are the truth-conditions of claims about reasons in ordinary language? Do these claims express propositions, give voice to attitudes, issue in commands, or perform some other function?
- What kind of linguistic mechanisms determine the meaning of sentences containing reason claims? What is the compositional semantics of natural language talk of reasons?

These latter questions will be the subject of this dissertation. Our questions here are thoroughly empirical. They are to be settled by careful attention to how ordinary speakers actually use the language of reasons. This type of investigation may well have consequences for philosophical theorizing about reasons and related matters—and I will argue that it does. But the influence should not go in the other direction. Our subject here is natural language semantics, not the semantics of the true normative theory of reasons revealed by some type of philosophical reflection. In short, my aim is not to say how we *should* talk about reasons but rather how we *do* talk about reasons.

I proceed as follows. Chapter 2 argues that claims about reasons are a type of modal language. Familiar modals like "ought", "might", and "must" describe how things stand with relevant bodies of information, e.g. "You must pay your taxes" describes the relevant laws as requiring you to pay your taxes. I show that reason claims describe relevant information in an analogous manner: *There is reason to believe that* ϕ describes the relevant *knowledge* as *counting in favor of* believing that ϕ. I then present a compositional semantics for reason-

talk that demonstrates how the semantics of reasons mirrors a plausible metaphysics of the reason-for relation.

Chapter 3 argues that the language of reasons pressures us to revise widely held views about the meaning of epistemic modals like "might". On the canonical semantics for epistemic modals, *it might be that* ϕ describes ϕ as compatible with the contextually relevant knowledge—hence the name "epistemic" modal. But I show that this theory faces a problem: if talk of reasons for belief and epistemic modal language both describe relevant knowledge, then we cannot explain why each type of language differs in how it embeds in complex sentences. I argue that the solution lies in abandoning the canonical semantics for epistemic modals: it is talk of reasons for belief that describes knowledge, while talk of what might be the case describes facts that may or may not be known by anyone. In addition, I argue that this semantic difference reveals a key feature of our pre-theoretical concept of a reason for belief. Many philosophers hold that reasons are simply facts—facts that count in favor of acts or attitudes. The reasons we possess are then understood as facts to which we enjoy some kind of epistemic access. I argue that this theory has things backwards. It is not as though the world contains reasons for belief that out there existing independently of human thought. Instead, a fact is intelligible as a reason for belief only in relation to a thinker who could easily come to be aware of this fact. This point about our concept of a reason for belief explains why talk of reasons for belief describes relevant knowledge: when we talk about reasons for belief, we talk about knowledge because this is what we take reasons for belief to *be*.

Chapter 4 uses the preceding investigations in semantics to draw out several consequences for epistemology. Once we see that "epistemic" modals have no special connection to knowledge, we are in a position to diagnose much of the confusion surrounding the formulation of fallibilism. It is often thought that fallibilism can be equivalently characterized as both (a) the view that knowledge is consistent with the possibility of error, and (b) the view that knowledge can be based on non-entailing evidence. But on my own semantics for epistemic modals, there is no equivalence between (a) and (b), since epistemic modals like "possible" do not describe states of evidence or any other epistemic state. In fact, I argue that epistemic modals have no place in the formulation of fallibilism at all. It is therefore a mistake

to rely on intuitions about epistemic modals to assess the truth of fallibilism, and I propose an alternative methodology for determining whether fallibilism is true.

Chapter 5 expands on the semantics for epistemic modals advanced in Chapter 3 by addressing the question of the semantics of probability modals. I argue that the semantics of probability modals is best understood in terms of a framework based on probability measures. I show that the most sophisticated non-probabilistic alternative faces important problems that have no analogue for a semantics based on probability measures. I conclude by sketching a novel version of the latter semantics that captures intuitive judgments about the logic of probability-talk even in the context of infinitely large domains of epistemic possibilities.

2.0 REASONS AS MODALS

2.1 INTRODUCTION

The aim of this chapter is to defend the following general claim: reason-talk is a type of modal language. More precisely, reason claims describe how things stand with relevant bodies of information, where the relevant information is determined by the same linguistic mechanisms that govern the interpretation of modals like "ought", "might", and "must". My aim here is to set out only the general shape of this semantics, focusing particularly on the language of so-called normative reasons, i.e. reasons that serve to justify acts and attitudes. I will argue that such reason claims exhibit key features that place them as part of modal language. And I will use this characterization to give a sketch of the compositional semantics of reasons. As we'll see, there turns out to be a connection between the *semantics* and *metaphysics* of reasons after all: I argue that the former actually reflects the latter.

2.2 BACKGROUND: THE LANGUAGE OF REASONS

I begin by setting out some data about the type and distribution of reason claims in natural language. This data will form the basis of the semantic investigation to follow.

Natural language talk of (normative) reasons comes in three main varieties:

Relational Reason Claims: *ϕ is (a) reason to ψ/reason for ψ-ing*

Reason Existentials: *there is (a) reason to ϕ/reason for ϕ-ing*

Possessive Reason Claims: *S has (a) reason to ϕ/reason for ϕ-ing*

Here are some examples of each that contain count noun uses of "reason":

(1) Anytime you get stopped, it is a reason to be afraid.[1]

(2) If there is a reason for hope, it is that the pundits' hostile account of conservatism has an element of truth.[2]

(3) Every sponsor has a reason for being here.[3]

Each of our three main varieties can also contain mass noun uses of "reason":

(4) The Lone Peak tram is reason enough to ski Big Sky (never mind the other 32 lifts).[4]

(5) There is reason to believe that not having a car isn't just a consequence of poverty—it's a barrier to escaping it.[5]

(6) For now, Miami has reason to cheer.[6]

Each variety is also gradable and participates in comparatives:

(7) Being laid off is an excellent reason to travel.[7]

(8) There's little reason to believe Brady won't be an elite player in 2015, but there's even less reason to believe he'll remain one in 2017, 2018 or whatever years lay down the road.[8]

(9) It's too expensive to live here unless you have a very good reason to do so.[9]

Finally, each type of reason claim readily embeds in a variety of constructions, such as questions, conditionals, and attitude reports:

(10) So, the rise in the sector made the indexes look more correlated. Is that a reason to throw international investing out the door? No.[10]

[1] "In metro Denver illegally, Arturo Garcia worked and lived in shadows," Tom McGhee, *Denver Post*, 2015.
[2] "Political Constitutionalism," Ramesh Ponnuru, *National Review*, 2014.
[3] "NASCAR's growth slows after 15 years in fast lane," Nate Ryan, *USA Today*, 2006.
[4] "Size Matters," Drew Pogge, *Skiing*, 2014.
[5] "Auto-Mobility," Margy Waller, *Washington Monthly*, 2005.
[6] "A Ballpark That May Be Louder Than the Fans," Michael Kimmelman, *New York Times*, 2012.
[7] Quoted in "Lights! Camera! Travel!; Travel Channel course gives aspiring journalists digital film training," Jayne Clark, *USA Today*, 2009.
[8] "NFL's Best QB Situations," Patrick Daugherty, *Rotoworld*, 2015.
[9] "You Can't Kill Mr. Goodbar," Alexander Nazaryan, *Newsweek Global*, 2015.
[10] Quoted in "Think Globally," James K. Glassman, *Washington Post*, 2004.

(11) If there is no pathology, if there's no pain and suffering, then there's no reason for medicine to be there.[11]

(12) The Rev. Kenneth Samuel believes he has a reason to be challenging incumbent Democratic Rep. Earnest "Coach" Williams in DeKalb County's District 89 primary.[12]

Let me turn now to my proposal for the semantics of this language.

2.3 THE MODAL MODEL

I will argue that the language of reasons is best understood as a species of modal language. This might appear to be a puzzling claim. Modal language is often understood as language concerning possibility and necessity. But it's not obvious that talk of reasons concerns either. It might also appear that talk of normative reasons is simply of the wrong linguistic category to count as part of modal language. "reason"—in the sense at issue here—is a noun, but we typically think of modal language as involving auxiliary verbs like "might" and "must", adjectives like "possible", and adverbs like "possibly". And even if we do sometimes employ modal nouns like "possibility", the noun "reason" does not itself refer to some possible state of affairs. What, then, could it mean to locate the language of reasons in the category of modal language?

The answer lies in a more expansive conception of modal language, perhaps more familiar to linguists than philosophers. On this conception, modal language is connected to the general phenomenon—sometimes known as "displacement"—in which we use language to talk about things other than the here and now. In particular, modal language allows us to talk about states of affairs that need not be real or actual.

Here is how Paul Portner puts the point:

> [M]odality is the linguistic phenomenon whereby grammar allows one to say things about, or on the basis of, situations which need not be real. Let's take an example: I say "You

[11] "The Greatest Gift: Genetic Engineering to Save a Child," John Stossel and Elizabeth Vargas, *ABC 2020*, 2007.

[12] "Voter ire a factor in races for state seats; Poll numbers reveal soaring disapproval with incumbents," Aaron Gould Sheinin, *Altanta Journal Constitution*, 2010.

should see a doctor." I am saying something about situations in which you see a doctor; in particular, I am saying that some of them are better than comparable situations in which you don't see a doctor. Notice that what I say can be useful and true even though you do not see a doctor. Thus, what I say concerns situations which need not be real (Portner (2009, 1)).

Now, it should be uncontroversial that reason claims are part of modal language in this broad sense. Suppose I say the following:

(13) There is reason to believe that a large asteroid is headed on a collision course with the Earth.

Intuitively, what I said can be true even if no one—including myself—actually forms the belief that an asteroid is going to hit the Earth. The sentence can also be true even if it turns out that the my data is misleading, and the asteroid will actually miss the Earth ever so slightly. The reason claim therefore allows us to talk about two states of affairs, neither of which are represented by the sentence as being actual: (i) the state of affairs in which someone believes that an asteroid will hit the Earth; (ii) the state of affairs in which an asteroid is on a course to hit the Earth.

However, if this were the only sense in which reason claims characterize potentially non-actual states of affairs, then my thesis of "reasons as modals" would be of little interest. Consider:

(14) I want to believe that you are telling the truth.

This sentence exhibits both of the modal features we just saw above in the case of reason claims. That is, (9) can be true even if no one actually believes you are telling the truth, and (9) can be true even if you are not actually telling the truth. What we see here is that the infinitive phrase *to believe that* ϕ itself carries a modal meaning: the phrase represents neither the belief that ϕ nor ϕ itself as being the case.

Let us refer to the broad category of modal language just discussed as the *general category of modality*. Contrast the latter with what I will call the *linguistic category of modality*. Modal language in the general sense characterizes potentially non-actual states of affairs.

Modal language in the more restricted, linguistic sense characterizes potentially non-actual states of affairs *in a particular way*—namely, by characterizing how things stand with relevant bodies of information.

For example, suppose I say:

(15) It might rain tonight.

On the received semantics for epistemic modals like "might", (10) describes rain as compatible with the contextually relevant knowledge—e.g. the speaker's knowledge, some group's knowledge, or the knowledge the speaker or group could easily come obtain. In other words, the modalized sentence allows us to talk about a potentially non-actual state of affairs—namely, rain tonight—by describing how things stand with some relevant body of information—namely, that the relevant knowledge is compatible with the content *it rains tonight*.

Other fragments of modal language are thought to work in exactly the same way. On Angelika Kratzer's canonical semantics (Kratzer (1977, 1981, 2012)), modals like "might", "must", "ought", and "should" characterize two different bodies of information: information determined by what she calls the *modal base* and the *ordering source*. Both are formally represented as functions from worlds to sets of propositions (i.e. sets of sets of worlds). For example, the modal base for the epistemic modal in (10) might be a function from a world w to the set of propositions known by the speaker at w. The intersection of these propositions forms what I will call the *domain* of the modal. The modal is then understood to quantify over the worlds in this domain, where f is a parameter representing the modal base:

> *might* ϕ is true at a context-modal base-world triple $<c,f,w>$ iff $\exists w' \in \bigcap f(w) : \phi$ is true at $<c,f,w'>$.

This semantics tells us that *might* ϕ is true at a point of evaluation just in case the modal domain at this point of evaluation contains a world at which ϕ is true. Put simply: the modal claim describes ϕ as compatible with the relevant information.

The ordering source plays a different role, one that is most frequently discussed in connection with modals expressings preferences, or characterizing how things stand with rules,

or goals. The ordering source functions to determine a set of propositions that induces a ranking of worlds. Consider a deontic modal claim:

(16) You must pay your taxes by April 15.

The ordering source here might be a function from a world to the propositions specifying the tax law in some relevant country at that world. These propositions will determine a ranking on worlds: worlds are more ideal the more propositions they make true. In our present example, worlds are more ideal the more laws are followed in these worlds. (16) then describes the best of these worlds as all being worlds at which you pay your taxes by April 15.[13]

Kratzer uses variations in the type of modal base and ordering source to give truth-conditions for a variety of different modals and different readings of one and the same modal. But what unites these semantics is the common core on which each modal characterizes how things stand with relevant bodies of information.

This information-describing semantics plays one further role that will be crucial for my discussion of reason claims. On the standard picture of modals like "ought", "might", and "must", the information they characterize can be specified in two different ways. As we saw with (10) and (16), modalized sentences often do not explicitly state the information that the sentence is characterizing. Here the information must somehow be gleaned from the larger context of use. However, there are also modalized sentences where the relevant information is made explicit. Devices for making the information explicit are known as "restrictors". Here is one example:

(17) According to weather report, it might rain tonight.

In (17), the modal "might" characterizes how things stand with the information contained in the weather report: the sentence is true just in case the information in the report is compatible with rain. It is also widely held that conditional antecedents can restrict modal domains:[14]

[13] I am passing over various complications. See Portner (2009) for a more detailed overview of Kratzer's semantics.

[14] Again, the canonical source here is Kratzer. See Kratzer (1991*a*).

(18) If you haven't heard back from the doctor yet, then the test was probably negative.

Here the modal "probably" describes the information that you haven't heard from the doctor as making it likely that the test result is negative.

In (18), the antecedent restricts the domain of an explicit modal in the consequent. But it is often thought that *all* conditionals have a modal in the consequent, either explicit, tacit, or sometimes even both. For example:

(19) If it's raining, then the streets are wet.

This sentence is typically thought to contain a tacit epistemic necessity modal with wide scope over the consequent. That is, the real structure of (19) is better represented as follows:

(20) If it's raining, then it must be that the streets are wet.

If all conditionals include such a modal in the consequent, then one might hold, as many do, that the function of conditional antecedents *just is* to restrict modal domains.[15]

With this background in place, I can state more precisely what I aim to establish in this chapter: reason claims characterize relevant bodies of information in the same manner as familiar modals. This is not just to say that reason claims fall within the general category of modality. I am advancing a substantive thesis that places the language of reasons within a very specific *linguistic* category.

In sum, I aim to defend what I will call the **Modal Model**:

> **The Modal Model**: Reason claims fall under the linguistic category of modality: they describe how things stand with relevant bodies of information, where the relevant information is determined by the same linguistic mechanisms that govern the interpretation of modals like "ought", "might", and "must".

[15] Again, see Kratzer (1991*a*).

2.4 GROUNDS FOR THE MODAL MODEL

I offer two main arguments for the **Modal Model**. On the first, I argue that reason claims engage with restrictors in the characteristic manner of modals. On the second, I argue that reason claims display the characteristic behavior of modals when no restrictor is present to settle the relevant body of information.

2.4.1 Reasons With Restrictors

Reason claims readily embed under restrictors. I focus here on restriction via conditional antecedents:

(21) If Jake cheats once, there's every reason to believe he'll do it again.

(22) If Jake cheats once, he'll probably do it again.

(23) If it's going to rain tonight, then you have no reason to bother washing the car now.

(24) If it's going to rain tonight, then you shouldn't bother washing the car now.

(25) If the drugs were planted, then the fact that they were found in Pete's car is not a reason to believe that he's a trafficker.[16]

(26) If the drugs were planted, then the fact that they were found in Pete's car does not make it likely that he's a trafficker.

Each pair of sentences appears to have roughly the same meaning: the consequent characterizes how things stand with the information contained in the antecedent. That is, (21) describes the information that a person has cheated once as constituting strong grounds for thinking they'll do it again, while (22) describes this information as making it likely that they'll cheat again. Similarly, (23) and (24) characterize the information that it's going to rain as undermining any grounds for washing the car now. (25) and (26) describe the relational claims in the consequent as being true relative to the body of information in the antecedent.

[16] Kotzen (ms.) discusses similar cases and uses them to motivate a modal semantics for relational reason claims. Henning (2014) offers a similar semantics motivated by data I discuss in the next section.

The point here is not just that these conditionals appear to have the same meaning. We can also give an argument for the claim that conditionals restrict the domains of reason claims that parallels the standard "no scope" argument for taking conditionals to restrict modal domains.[17]

Suppose the conditional is a two-place operator, either the material conditional or the strict conditional. This operator must either scope above a reason claim or below a reason claim. But neither option yields the right predictions.

Recall (21):

(21) If Jake cheats once, there's every reason to believe he'll do it again.

Suppose the conditional is a material conditional with wide scope over the reason claim. This semantics wrongly predicts that (21) is true if cheating once has no correlation to cheating again but Jake never cheats. Suppose instead that the reason claim scopes over the conditional, i.e. the logical form (21) is really represented by (27):

(27) There is every reason to believe that if Jake cheats once, he'll do it again.

This sentence is mistakenly predicted to be true if cheating once has no correlation to cheating again but there is every reason to believe that Jake will never cheat.

Suppose instead that the conditional is a strict conditional with wide scope over the reason claim. (21) now claims that every world (or every epistemically accessible world) in which Jake cheats once is a world where there is reason to believe that Jake will cheat again. But this seems too strong. Just consider a world where Jakes cheats once but the speaker is unaware of this fact. There is a reading of the reason claim where it is true in this scenario that there is no reason to believe Jake will cheat again because there is no reason to believe Jake has cheated in the first place. Finally, suppose the strict conditional scopes below the reason claim, as in (27). The resulting truth-conditions are again too strong: there might be no reason at all for thinking that *every* epistemically accessible world where Jakes cheats once is a world where he cheats again.

All of these problems are avoided if we take reason claims themselves to be inherently relational: reason claims tell us how things stand with some relevant body of information

[17] See von Fintel and Heim (ms.) for an example of the latter.

that can be provided explicitly by a conditional antecedent. On this semantics, there is no two-place conditional operator, there is only a relational reason claim, one of whose relata—here a body of information—can be given by a conditional antecedent. This semantics bypasses the problems of the material conditional analysis since the falsity of the antecedent no longer verifies the entire conditional. The truth of the conditional turns on the *relation* between the information contained in the antecedent and the proposition in the scope of the reason claim, e.g. the truth of (21) turns on whether the fact that Jake has cheated once counts in favor of thinking he'll cheat again. The restrictor semantics bypasses the problems of the strict conditional analysis since the reason claim in (21) is evaluated only *given* the information contained in the antecedent. Nothing follows about what there is reason to believe in worlds where the antecedent is merely true. The restrictor semantics thus allows for the possibility of scenarios where Jakes cheats once but there is no reason to believe he'll cheat again. These will just be scenarios in which the reason claim is not evaluated with respect to the information that Jake has cheated once. Such scenarios are entirely consistent with the truth of (21), as the latter explicitly supplies the information that Jake has cheated as the information in light of which the reason claim is assessed.

These considerations are suggestive but not decisive. I suspect that many will object that the restrictors in our examples with reasons are actually restricting a tacit modal in the consequent, such as "must" or "will". In other words, perhaps the examples do provide evidence for the restrictor analysis of conditionals. But the examples just involve the restriction of modal domains, not the domains of reason claims themselves.

However, it's doubtful that these strategies can explain the data in every case. Take (21) and (23). Both are naturally construed as holding *right now*, given the truth of the antecedent. That is, if Jake cheats, then given this fact, there is *right now* every reason to believe he'll do it again. The sentence has no reading on which it specifies how things stand at some interdeterminate future time. It's therefore implausible that the sentence contains a tacit modal "will" that scopes over the consequent.

There is also data that speaks against the presence of a tacit epistemic necessity modal, "must", but it will take more work to set out these cases. The data here concerns the role of reason claims in arguments that appear to invalidate classically valid inference patterns.

As background to these cases, consider the so-called Miner's Paradox:

> Ten miners are trapped either in shaft A or in shaft B, but we do not know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block one shaft, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed (Kolodny and MacFarlane (2010, 115)).

Kolodny and MacFarlane (2010) note that the following all seem true in the above scenario:

P1. We ought to block neither shaft.

P2. If the miners are in shaft A, we ought to block shaft A.

P3. If the miners are in shaft B, we ought to block shaft B.

And the following is true by hypothesis:

P4. Either the miners are in shaft A, or the miners are in shaft B.

From P1–P4 we can derive the following conclusion by Disjunction Elimination, Modus Ponens, and Disjunction Introduction:

C. Either we ought to block shaft A, or we ought to block shaft B.

But C is inconsistent with P1: if we ought to block one of the shafts, then it's false that we ought to block neither shaft.

Kolodny and MacFarlane argue that the solution to this paradox lies in rejecting Modus Ponens: this rule of inference is not generally valid and the above constitutes a counterexample.

Consider P2. Kolodny and MacFarlane claim that the antecedent directly restricts the domain of the modal in the consequent. The modal thus describes blocking shaft A as the best course of action given the information that the miners are in shaft A.[18] Now suppose that as a matter of fact, the miners are in shaft A. It does not follow that we ought to block

[18]This proposal requires a non-trivial departure from Kratzer's semantics for deontic modals discussed above. See Kolodny and MacFarlane (2010) and Cariani, Kaufmann and Kaufmann (2013) for discussion.

shaft A, since this modal claim does not necessarily characterize what we should do in light of the information about the miners' actual location. The conditional antecedent forces the modal in the consequent of P2 to characterize this further information. But if the modal is not in the scope of this restrictor, then it may well characterize some completely different body of information—one that need not make blocking shaft A the best course of action. This is how P2 and its antecedent can both be true without the consequent itself necessarily being true.[19]

We can construct a similar paradox involving reason claims.[20] Suppose again that the miners are either all in shaft A or all in shaft B, and we don't know which. But suppose instead that we have enough sandbags to block both shafts—although the sandbags are very heavy, so we'd prefer not to move them unless it is absolutely necessary.

In this scenario, the following seems true:

P1. There is reason to block shaft A, and there is reason to block shaft B.

Were we to block only one shaft or fail to block either, we would be properly criticized on the grounds that there was every reason to block both. This is true regardless of whether we happened to block the right shaft and save all of the miners.

The following also seem true:

P2. If the miners are in shaft A, then there's no reason to block shaft B.

P3. If the miners are in shaft B, then there's no reason to block shaft A.

Both claims seem perfectly obvious in the context of deliberation about what to do.

Finally, recall our original assumption:

P4. Either the miners are in shaft A, or the miners are in shaft B.

From P1–P4 we can derive the following conclusion by Disjunction Elimination, Modus Ponens, and Disjunction Introduction:

[19]See Kolodny and MacFarlane (2010) for discussion of various objections one might raise against this proposal.

[20]The case to follow is different from Kolodny and MacFarlane's in several respects, as will become clear below. I have chosen to modify the case so as to avoid using the awkward phrasing of "most reason" in place of "ought". See Henning (2014) for discussion of a case involving the "most reason" locution.

C. Either there's no reason to block shaft A, or there's no reason to block shaft B.

However, C is inconsistent with P1: it is a basic fact about the logic of reasons that *there's no reason to* ϕ entails the negation of *there is reason to* ϕ. So if one of the shafts is such that there is no reason to block it, then it cannot be the case that there is reason to block shaft A *and* reason to block shaft B.

Again, this data is not decisive. But it constitutes a strong case for thinking that reason claims have domains that can be restricted directly by conditional antecedents. Consider:

(i) If the miners are in shaft A, then there's no reason to block shaft B.

(ii) The miners are in shaft A.

(iii) There's no reason to block shaft B.

Suppose the antecedent of (i) directly restricts the domain of the reason claim in the consequent. That is, the reason claim describes the information that the miners are shaft A as undermining any grounds for blocking shaft B. Now suppose that the miners are in shaft A. It does not follow that there is no reason to block shaft B, since the latter reason claim is not within the scope of a restrictor. That is, (iii) may well be characterizing how things stand with some different body of information that *does* constitute reason to block shaft B, and so the reason claim may well be false. Notice a crucial feature of this case: if there is a tacit epistemic necessity modal in the consequent of (i) that is restricted by the antecedent, then *the inference to (iii) will be valid.*[21] But it isn't.[22] So the antecedent restricts the reason claim directly.

Willer (2012) pursues an alternative strategy for resolving the original Miner's Paradox. He presents a semantics on which Modus Ponens is valid, but Disjunction Elimination is not. However, his semantics still invalidates Modus Tollens. Consider, then, how one might employ his semantics to resolve our puzzle about reasons. Take the following inference:

[21]Suppose there is a tacit epistemic necessity modal in the consequent restricted by the antecedent. (i) will then express the following proposition: every epistemically accessible world in which the miners are in shaft A is a world in which there is no reason to block shaft B. Hence, if we're in a world where the miners are in shaft A, then we must be in a world where there's no reason to block shaft B (I make the standard assumption that the actual world is always epistemically accessible).

[22]Note: one must take care not to confuse the above inference with an inference in which one draws the conclusion on the basis of *knowledge* of the premises. The latter is perfectly valid even if Modus Ponens is not. See Kolodny and MacFarlane (2010) for discussion.

(a) If the miners are in shaft A, then there's no reason to block shaft B.

(b) There is reason to block shaft B.

(c) The miners are not in shaft A.

(a) and (b) seem true in our above scenario, even though the conclusion (c) does not intuitively follow. These judgments support a semantics like Willer's that invalidates Modus Tollens. However, notice again what would happen if the antecedent of (a) were restricting the domain of a tacit epistemic necessity modal with wide scope over the consequent: the inference to (c) would be vaild.[23] But it isn't. So in general: if one wishes to resolve our paradox about reasons by denying the validity of some classically valid inference pattern, then one must take conditional antecedents to restrict the domain of reason claims directly.

2.4.2 Reasons Without Restrictors

As we noted in §2.3, modals often appear without restrictors that explicitly specific which information is relevant. Such modals are known as "bare modals". Data about bare modals has been the topic of much recent discussion, as it presents various foundational challenges to a simple, contextualist picture of the semantics of modals.[24] I will not take sides in this dispute. I simply aim to show that "bare" reason claims give rise to a parallel set of data and that this motivates a structurally parallel semantics.

The puzzle begins from the observation that epistemic modal talk does not behave as one would expect if it simply describes what the speaker knows. Here is an early example from Huw Price:

> If I disagree with your claim that it is probably going to snow, I am not disagreeing that given *your* evidence it is likely that this is so; but indicating what follows from *my*

[23] Suppose there is a tacit epistemic necessity modal in the consequent restricted by the antecedent. (i) will then express the following proposition: every epistemically accessible world in which the miners are in shaft A is a world in which there is no reason to block shaft B. Hence, if we're in a world where there is reason to block shaft B, then we cannot be in a world where the miners are in shaft A.

[24] See, for instance, Egan, Hawthorne and Weatherson (2005), Egan (2007), Dowell (2011), von Fintel and Gillies (2011), Khoo (2015), MacFarlane (2011, 2014), Moss (2013, 2015), Rothschild (2012), Swanson (2006, 2011, 2016), Willer (2013), Yalcin (2007, 2010, 2011, 2012*a*,*b*), and Yanovich (2014).

> evidence. Indeed, I might *agree* that it is probably going to snow and yet think it false that this follows from your evidence (Price (1983, 404); his emphasis).

The initial problem is straightforward. If "it's probably going to snow" just means "My evidence suggests it's likely to snow", then it's not clear why a more informed listener might take herself to disagree with a speaker's probability claim. After all, she need not disagree that the *speaker's* evidence suggests that it's likely to snow. Conversely, a more informed listener might take herself to agree with what the speaker said but disagree that the speaker's evidence suggests that it's likely to snow. A natural suggestion is to take the speaker's claim to instead characterize how things stand with the evidence available to some larger group. But it then becomes unclear why the speaker was warranted in making her assertion in the first place.

My aim is not to canvass the various attempts to solve this puzzle. I only aim to show that reason claims present the very same challenge. Suppose you have a friend who believes he can predict the weather by harnessing the power of crystal skulls. He comes to you one day and asserts the following:

(28) There is good reason to believe that a major storm is headed this way.

If you've already checked the weather report, and meteorologists are also predicting a storm, then there's clearly some sense in which you and your friend agree. And this can be so regardless of the fact that you take your friend's "evidence" to be completely irrelevant to whether a storm is going to occur.

Here is a parallel case involving disagreement. Suppose Jones is a fine, up-standing citizen who happens to live in a town where a murder has just occurred. Someone might remark:

(29) There's no reason to believe that Jones was involved.

But suppose the police have just learned that Jones's blood was found on the murder weapon. If they overhear (29), they might reply:

(30) Actually, there is: we just found his blood on the murder weapon.

This use of "actually" typically functions to signal disagreement.[25] But of course, the police might agree that the initial speaker had no evidence for thinking Jones was involved.[26]

So far I've been focusing on intuitions about bare reason existentials. But we can elicit the same type of intuitions with bare relational reason claims and bare possessive reason claims. For example, suppose the police find large quantities of narcotics in Pete's vehicle. A detective might assert the following:

(31) We have every reason to believe that Pete is a drug trafficker.

But if you've just learned that the drugs were planted, you can reply:

(32) No you don't—the drugs were planted!

Similarly, if someone says:

(33) The fact that drugs were found in Pete's car is a reason to believe he's a trafficker.

You can reply:

(34) No it isn't—the drugs were planted!

None of this should be surprising if reasons claims characterize relevant bodies of information. If the relevant information is not specified explicitly, then it must be settled in some fashion by the context of use or perhaps left open as a further parameter of evaluation. It is not straightforward how any of this works in the case of familiar modals. But if we recognize a

[25] Compare:

A: I'm not hungry.
B: # Actually, I am.

B's reply sounds bizarre if it's read with the kind of intonation we naturally place on the phrase given in the main text: "Actually, there is." In particular, one should make sure not to place focal stress on the word "I". A simple explanation of the infelicity of B's reply is that "actually" signals disagreement, but what B said does not conflict with what A said.

[26] Agreement and disagreement intuitions are not entirely straightforward. Some speakers judge the purported cases of disagreement as instances of two people talking past one another. However, the same goes for intuitions about eavesdropping scenarios concerning epistemic modals. Some theorists even hold that the relevant datum is just that two readings are available, one involving disagreement and one not (see, for example, von Fintel and Gillies (2011)). The data about reason-talk plausibly involves the very same kind of variability, which provides further evidence for the parallelism between reason-talk and modal-talk.

parallel semantics for reasons, we are a least in a position to give these problems a unified resolution.[27]

To sum up: We've seen that reason claims give rise to the characteristic behavior of modal claims in a variety of different cases, both with and without restrictors. The types of cases at issue have been the focus of extensive discussion in the recent literature on modals, with numerous authors claiming that these cases reveal distinctive features of the semantics of modals. The **Modal Model** offers a simple explanation for all of these parallels: reason claims characterize relevant bodies of information in the very same manner as modals.

2.5 COMPOSITIONAL SEMANTICS

I conclude with a brief sketch of the compositional semantics of reasons. My aim here is to set out the general shape that such a semantics should take and to show that the **Modal Model** can be made precise with the tools of formal semantics.

Now, I have emphasized from the start of this dissertation that we must distinguish between the *semantics* of reasons and the more familiar investigation into *metaphysics* of reasons that has been the main focus of philosophers in the past several decades. However, it turns out that the latter investigation has identified precisely the sort structure we need to give a semantics for ordinary language talk of reasons.

T.M. Scanlon's account of the reason-for relation will provide the framework for my semantics. Here is how he characterizes the reason-for relation:

> "is a reason for" is a four-place relation, $R(p, x, c, a)$, holding between a fact p, an agent x, a set of conditions c, and an action or attitude a. This is the relation that holds just in case p is a reason for a person x in situation c to do or hold a (Scanlon (2014, 31)).

Let me take each relatum in turn, first explaining its role in a semantics for reasons, and then turning later to give a fully compositional semantics for reason claims.

[27] See Chapter 3 for further discussion.

Start with the relatum p. p is itself the reason. As Scanlon puts it elsewhere, reasons are things that count in favor of acts or attitudes.[28] So p is the thing that counts in favor of these acts or attitudes. The other relata specify what p counts in favor of, for whom, and in what circumstances.

Now, Scanlon takes p to simply be a fact. I will not adopt this assumption in my semantics. However, my proposal is similar: the things that we talk about when we talk about reasons are propositions, sets of propositions, or mereological sums of propositions. Some of these propositions may be true at the world of evaluation, and so they are facts, as Scanlon says. But I wish to leave open the possibility that some of our talk of reasons may refer to falsehoods, on the one hand, and knowledge on the other. That is, natural language talk of reasons is not simply talk of facts.[29]

Continuing, on Scanlon's metaphysics, a reason is always a reason for something, for someone, and in some circumstances. Each of the preceding plays a separate role in my semantics. Consider a relational reason claim:

- *ϕ is a reason to believe that ψ.*

If a sentence of this form is true, then ϕ must be a reason: in a simple case, it will be a proposition. But it can't just be any proposition. It must be a proposition that counts as a reason *to believe that ψ*.

Let's take apart the meaning of the infinitive phrase *to believe that ψ*. The attitude verb "believe" plays the role of Scanlon's relatum a: the verb specifies what general type of act or attitude the reason is a reason for. Scanlon's person x appears covertly. Infinitive phrases like *to believe that ψ* are standardly thought to contain a hidden pronoun, known as PRO. The referent of this pronoun is the agent who does the believing—here, the agent for whom the reason is a reason to believe that ψ. It is often thought that a reason for belief is a reason for anyone to hold the given belief.[30] So we should allow that PRO can refer to some generic subject x. Claims about reasons for action work differently: the very same proposition will

[28] See Scanlon (1998, Chapter 1).

[29] See Chapter 3 for arguments that talk of reasons for belief describes information connected to thinkers' epistemic states. See Comesaña and McGrath (2014) and Schroeder (2008) for arguments in favor of non-factive reasons.

[30] See Schroeder (2008).

often count as a reason for one person to perform an action but not another. For example, if one person likes dancing but another does not, then the fact that a party will involve dancing is a reason for the first person to go but not the second.[31] We should allow, then, that the referent of PRO can pick out a particular subject for whom dancing is a reason to go to the party.

I wish to remain neutral on how exactly the referent of PRO is determined in a given case. One might hold that PRO's referent is settled by the context of use. But one might also allow that PRO's value is fixed grammatically—i.e. "controlled"—by some other element of the discourse. There may even be cases where PRO is bound by a quantifier, as in (35):

(35) Everyone has a reason to go to the party.

This sentence describes the party as being such that there is a reason for each individual person to attend, i.e. the party contains something for everyone.

Now, as Scanlon emphasizes, a proposition may count as such a reason in one set of circumstances but not another even when we hold fixed who the reason is a reason for. We've already seen one example of this phenomenon when we discussed relational reason claims embedded under conditionals:

(36) If the drugs were planted, then the fact that they were found in Pete's car is not a reason to believe that he's a trafficker.

Ordinarily, the fact that drugs were found in Pete's car is a reason to believe that he's a drug trafficker. But not so in circumstances where the drugs were planted. The role of the conditional antecedent in (36) is to make explicit what the relevant circumstances are for the evaluation of the reason claim in the consequent. That is, the conditional antecedent shifts the relevant circumstances so that they include the information in the antecedent. This is why (36) sounds true. This type of circumstance-relativity is the first element of the modal meaning of reason claims: relevant circumstances constitute a body of information, so reason claims characterize bodies of information along at least one dimension. We have, then, the following truth-conditions for relational reason claims:

[31] This example comes from Schroeder (2008).

(a) ϕ *is a reason to* ψ is true at a world-circumstance pair $<w,c>$ iff ϕ is a reason for $[\![\text{PRO}]\!]$ to ψ at w in c.

(b) ϕ *is reason to* ψ is true at $<w,c>$ iff ϕ is reason for $[\![\text{PRO}]\!]$ to ψ at w in c.

(a) gives the truth-conditions for a count noun relational reason claim: *is a reason to* ψ determines a property of propositions that is predicated of ϕ. (b) gives the truth-conditions for a mass noun relational reason claim: *is reason to* ψ determines a property of either some set of propositions, the intersection of this set, or perhaps the mereological sum of the members of this set. The choice here will be determined by the semantics of mass nouns in general.

Now, so far I've only discussed the semantics of relational reason claims, wherein the reason—ϕ—is specified explcitly even if no restrictor is present. But bare reason existentials and bare possessive reason claims fail to state which proposition counts as the reason:

- *there is (a) reason to* ϕ
- *S has (a) reason to* ϕ

I propose that in cases like these, the reason is given by a modal base f. The value of this function at a world will be a set of propositions. We can take the intersection of this set to provide the proposition that constitutes the reason (again, in the case of mass noun reason claims, we might instead treat the entire set or perhaps the mereological sum of the members of the set as constituting *reason*).

This proposal easily accommodates the phenomenon whereby conditional antecedents restrict the domain of reason claims: the domain is simply given by the intersection of the value of the modal base at the world of evaluation, and conditional antecedents shift the parameter f to the include the information in the antecedent, as on Kratzer's original semantics. This proposal also accommodates the variability in reason claims without restrictors: there may be no unique value for f in a typical context of use, which may give rise to the puzzling intuitions discussed in §2.4.2.[32] The modal base f is thus the second modal element in our semantics for reasons.

[32]See Chapter 3 for the details of this explanation.

Let me say more about the particular role that f plays in our semantics. I will take reason existentials to contain quantifiers that directly engage with the modal base f:

(a) *there is a reason to* ϕ is true at $<w,f,c>$ iff $\exists p \in f(w) : p$ is a reason for $[\![\text{PRO}]\!]$ to ϕ at w in c.

(b) *there is reason to* ϕ is true at $<w,f,c>$ iff $\exists p = \bigcap f(w) : p$ is a reason for $[\![\text{PRO}]\!]$ to ϕ at w in c.

(a) corresponds to the natural reading of the count noun reason existential on which it characterizes how things stand with at least one piece of the relevant information. (b) corresponds to the natural reading of the mass noun reason existential on which it characterizes how things stand with the entire body of relevant information.

Possessive reason claims can receive identical readings when the subject S is the speaker. For example, there is little difference between (37) and (38):

(37) There is reason to believe that Jones is the murderer.

(38) I have reason to believe that Jones is the murderer.

However, when S is someone other than the speaker, it's natural to hear the sentence as characterizing how things stand with S's information:

(39) John has reason to believe that I'm the murderer.

This sentence has a reading on which it is true even if the speaker knows that she is not a murderer. It is a further question how phrases like "John has" yield this reading. I leave this as an issue for future research.

To recap: I've offered a semantics for reasons that employs three key elements—the hidden pronoun PRO, the modal base parameter f, and the circumstance parameter c. It is important to see why all three elements are needed. It is true that a single, centered world will contain exactly the same information as our set of three elements. But the problem with a centered world-based semantics is that we need to be able to shift each element independently. As we've seen above, conditionals sometimes change the body of information that constitutes the reason (see (40) below), but they can also leave the reason unchanged

while varying the set of circumstances in which the reason is evaluated as a reason (see (41) below).

(40) If there are drugs found in Pete's car, then there is reason to believe that he is a trafficker.

(41) If the drugs were planted, then the fact that they were found in Pete's car is not a reason to believe that he's a trafficker.

We've also seen that talk of what someone has reason to believe/do can shift either the value of PRO or the modal base parameter. For example, suppose you know that an assassin is about to target Bill in his office but Bill is unaware of this. The following seems true:

(42) Bill has a reason to leave the building even though he doesn't realize it.

Here the modal base contains the information known to the speaker; the modal base is not restricted to Bill's information. The locution "Bill has" functions to change the value of PRO: *Bill* is the one for whom the information about the assassin is a reason to leave the building. A centered world-based semantics makes the wrong predictions about this case. Such a semantics guarantees that the relevant information and the agent for whom that information is a reason are always given by one single individual—namely, the center in the relevant centered world. But (42) shows that the relevant information state can sometimes be tied to one individual, while the relevant agent is a completely different person.

Now, what I've set out so far is only a general sketch of the semantics of reasons. I have not specified that in virtue of which any particular proposition counts as a reason for an act or attitude. How—or even whether—to give a more informative semantics is a difficult question that I cannot discuss in detail here. Scanlon takes the reason-for relation to be primitive, and so he would deny that a more informative semantics is possible. Others have argued in favor of a reductive account of this relation.[33] There is also a middle ground position wherein one supplies some additional structure to the relation without offering a completely reductive semantics. But these are topics for another occasion.

Still, there is one final item on our agenda for this chapter. I've set out a proposal for the general truth-conditions of the language of reasons, but it remains to be seen how exactly

[33] See Finlay (2014).

this proposal can be implemented compositionally. I will close with some brief remarks on how one might carry out this task.

The output of the composition process is already settled by the truth-conditions given above: a reason claim determines a function from world-modal base-circumstance triples to truth-values.[34] Let us work backwards to determine the denotations of the elements of the reason claim. I said above that in the simplest case, reasons will be propositions. So a phrase like *is* (*a*) *reason to* ϕ should denote a property of propositions: that is, a function from propositions to functions from world-modal base-circumstance triples to truth-values. The general structure of this function is implicit in the truth-conditions I gave above for the various types of reason claims. The particular function denoted by *is* (*a*) *reason to* ϕ will be settled by the referent of the hidden pronoun PRO and the infinitive phrase *to* ϕ. However, it is a vexed question what denotation we should assign to the word *to* and the resulting infinitive phrase. For now, we will have to rely on a placeholder: let us take the phrase PRO *to* ϕ to denote a proposition—i.e. a function from worlds to truth-values. In particular, we can take the phrase to denote the function from a world w that returns truth iff the referent of PRO ϕ's in w. The end result is the following denotation for the word "reason": the semantic value of this word is a function from propositions to functions from propositions to functions from world-modal base-circumstance triples to truth-values.

[34] I leave open the role of the context coordinate of the index. Perhaps it settles the value of the modal base parameter. Or perhaps the value of this parameter is left open, to be settled by the context of assessment. See Chapter 3 for discussion.

3.0 THE SEMANTICS OF EPISTEMIC REASONS AND EPISTEMIC MODALS

3.1 INTRODUCTION

Suppose the head of NASA calls a press conference and announces the following:

(1) There is reason to believe that a large asteroid is headed on a collision course with the Earth.

The public's behavior will change dramatically. Why? A simple answer is that (1) conveys a non-trivial degree of confidence in an asteroid's hitting the Earth. But of course, (1) does not convey *certainty* or *near certainty* that the asteroid will hit. The claim conveys a far weaker but still significant degree of confidence.

Sentence (1) has the following general form:

There is (*every*/*good*/*little*/*no*) *reason to believe that* ϕ

Call sentences with this form "epistemic reason claims", or "reason claims" for short. As we just saw with (1), reason claims are important. Their significance is even enshrined as a matter of U.S. law, with the Supreme Court holding that the police may lawfully enter a suspect's home only "when there is reason to believe the suspect is within."[1] It's hard to

[1]See *Payton v. New York*, 445 U.S. 573 (1980) at p. 603. See Rabasca (2012) for discussion of subsequent disputes over how to interpret the Court's "reason to believe" standard. Philosophers and lawyers aren't the only ones who talk about what there is reason to believe. The Corpus of Contemporary American English yields 864 unique results for *There 's*/*is* (*every*/*good*/*little*/*no*) *reason to believe*/*think*. And the phrase *reason to believe* is in fact the most frequently occurring of all *reason to* [verb] constructions. Relational reason claims—i.e. claims of the form ϕ *is a reason to believe that* ψ—turn out to be extremely rare, even though these constructions are probably the most commonly discussed epistemic reason claims among philosophers. The corpus lists 9 results for *is* (*not*) *a* (*good*) *reason to believe/think*, only 3 of which are

understand the importance we place on reason claims if they convey no degree of confidence in a proposition's truth. And yet the degree of confidence they do convey falls well short of certainty. The upshot is that reason claims are part of what has become known as the "language of subjective uncertainty"(Swanson, 2011).

Debates about the semantics of this language have occupied center stage in much recent literature. Numerous authors claim that traditional semantics fails to capture the communication of subjective uncertainty, and proposals for accommodating such communication range from expressivism, to relativism, to non-standard versions of contextualism.[2]

To date, this research has focused almost exclusively on the semantics of epistemic modals, or what I will call "modal claims" for short—e.g. sentences of the form *it is likely that* ϕ, *it is probable that* ϕ, or *it might be that* ϕ. But as we've just seen: reason claims convey uncertainty as well. However, once we locate reason claims within the language of subjective uncertainty, questions immediately arise: What is the relation between the meaning of reason claims and the meaning of modal claims? What parallels, if any, exist between the semantics of these two types of expressions?

In this paper, I argue that there are strong ties between reason claims and modal claims that raise a puzzle about the semantics of epistemic reasons. And I argue that this puzzle ultimately transforms our understanding of the meaning of epistemic modals.

The puzzle stems from the fact that reason claims exhibit the characteristic behavior of modal claims in so-called eavesdropping scenarios—that is, scenarios in which an utterance is assessed by a third party with different information than the one who originally made the claim. The eavesdropping data about modal claims has been a driving force behind many of the recent proposals about the distinctive semantics of epistemic modals.[3] The parallel data about reason claims suggests they will have some kind of structurally similar semantics.

---

relational reason claims (corpus last accessed June 25, 2017). Finally, the language of epistemic reasons is not unique to English. Native speakers of Cantonese, Dutch, German, Italian, and Spanish confirm that these languages contain roughly synonymous expressions.

[2]See, for instance, Egan, Hawthorne and Weatherson (2005), Egan (2007), von Fintel and Gillies (2011), MacFarlane (2011, 2014), Moss (2013, 2015), Rothschild (2012), Swanson (2006, 2011, 2016), Willer (2013), and Yalcin (2007, 2010, 2011, 2012*a*,*b*).

[3]See, for instance, Dowell (2011), Egan, Hawthorne and Weatherson (2005), Egan (2007), von Fintel and Gillies (2011), Hawthorne (2007), Khoo (2015), MacFarlane (2011, 2014), Willer (2013), Yalcin (2011), and Yanovich (2014).

However, it turns out that reason claims "embed" differently than modal claims: each type of expression behaves differently when placed in certain complex sentences. And this difference in embedding behavior reveals that *none* of the leading semantics for modal claims can be extended to reason claims.

Put simply, then, the puzzle is this: What is the semantics of epistemic reasons? The similarities with modal claims call out for a shared semantics. But the differences appear to show that a shared semantics cannot be had. I argue that the solution to this puzzle lies in re-thinking the semantics of epistemic modals. On the canonical semantics for epistemic modals, tracing back at least to G.E. Moore (1962), modal claims describe how things stand with bodies of knowledge. However, I will argue that it is in fact reason claims, not modal claims, that describe knowledge. The function of modal claims is instead to describe how things stand with the relevant *truths*, where there is no requirement that these truths be known or otherwise related to anyone's epistemic state.

This is a surprising result, for it means that the semantics of epistemic reasons actually pressures us to revise widely held views about the semantics of epistemic modals. But the payoff is clear: the theory I offer enables us to give a unified account of the language of subjective uncertainty that allows for divergence among the particular fragments of this language. On my view, reason and modal claims share an underlying semantics on which both expressions characterize relevant bodies of information, and this structural similarity explains the shared eavesdropping data. Nevertheless, each expression differs in which *type* of information it characterizes—knowledge vs. mere truths—and this difference explains the divergence in embedding data. I close by reflecting on the larger consequences of this theory. I argue that the semantics of epistemic reasons ultimately helps us understand our *concept* of an epistemic reason.

3.2 THE DATA

I begin by setting out two new sets of data concerning the relation between reason claims and modal claims. As we'll see, the parallels in eavesdropping data suggest that these expressions

have some type of shared semantics, but the differences in embedding data rule out extending the leading semantics for modal claims to reason claims.

3.2.1 Eavesdropping Data

Suppose Pete's vehicle has been stopped and found to contain large quantities of narcotics. He is arrested and jailed pending a hearing that will determine his bail. At the hearing, the District Attorney reviews the facts of Pete's arrest and asserts the following sentence:

(2) There is good reason to believe that Pete is a drug trafficker.

There are two ways the conversation can proceed from here.

> **Scenario #1:** Detective Barnes conducts a raid on Pete's home and discovers additional narcotics. She arrives at the hearing and presents additional testimony on behalf of the prosecution. Barnes remarks, "I agree with the DA." Or: "What the DA said is true." In general terms, Barnes *accepts* the DA's assertion.
>
> **Scenario #2:** Barnes has just discovered that, unbeknownst to the DA, the drugs found in Pete's vehicle were planted by a corrupt cop with a personal vendetta against Pete. Barnes rushes to the hearing only to arrive when the DA asserts (2). Barnes replies, "No there isn't—the drugs were planted!" Or: "That's false!" In general terms, Barnes *rejects* the DA's assertion.

Any adequate semantics for reason claims should enable us to explain why the conversation can proceed in these two ways. However, this constraint turns out to be surprisingly difficult to satisfy.

Consider a simple, contextualist semantics for reason claims: these expressions are context-sensitive descriptions of how things stand with the speaker's knowledge, on par with sentences like "All of the evidence I'm aware of suggests that P." This semantics explains why reason claims are important: their significance lies in the significance of determining how things stand with what one knows. And the semantics fits with the familiar theoretical characterization of reasons for belief as items of knowledge.[4]

[4]See, for instance, Hawthorne and Stanley (2008), as well as Williamson's (2000) related discussion of the nature of evidence.

However, this semantics makes it very mysterious why **Scenario #2** is a perfectly appropriate continuation of the above conversation. If the semantics is correct, then by uttering (2), the DA asserts that everything *she* is aware of suggests that Pete is a trafficker. And surely this proposition is true. Even if Barnes learns the drugs were planted, she should not dispute that *what the DA knows* suggests that Pete is a trafficker.[5] So if the DA is simply describing the state of her own body of knowledge, why is Barnes entitled to reject the DA's assertion?

One response is to modify our semantics so that reason claims describe how things stand with the relevant *group's* knowledge, where the group extends beyond the speaker. If Barnes counts as part of the relevant group, then in **Scenario #2**, what the DA said is false, since the group's knowledge does not suggest that Pete is a trafficker. No wonder Barnes is entitled to reject the DA's assertion.

The trouble is that it now becomes mysterious why the DA was warranted in making her assertion in the first place. The DA is not in a position to know how things stand with the information available to Barnes, so how is the DA entitled to pronounce upon the bearing of their pooled body of knowledge? One might object that the DA can make reasonable assumptions about what Barnes knows. But note that *anyone* who learned the drugs were planted could have appropriately rejected the DA's assertion. And surely the DA is not warranted in making assumptions about how things stand with *everyone's* pooled knowledge.

In general, the contextualist faces a dilemma: if reason claims describe the speaker's knowledge, these claims are easy to assert but too hard to dispute; alternatively, if reason claims describe knowledge beyond the speaker's grasp, these claims are easy to dispute but too hard to assert.

The same type of dilemma has attracted a great deal of attention in the recent literature on epistemic modals.[6] Talk of what might be the case is often thought to describe how things stand with the speaker's knowledge: *it might be that* ϕ is true at a context just in

[5] If you are inclined to doubt this point, ask yourself: What can the DA say to defend her actions if it later comes out that the drugs were planted? Answer: she can state, truthfully, that everything *she* knew suggested that Pete was a trafficker.

[6] See the papers cited in n. 3.

case what the speaker knows is compatible with ϕ.[7] However, if I say that the keys might be under the doormat, you can reject my assertion by saying "No, I already checked there." And you can reject my assertion while being perfectly aware of the fact that everything *I* knew was consistent with the keys being under the doormat. As before, one might try to accommodate this data by taking my assertion to describe our combined knowledge. But again, this proposal makes it mysterious why I'm warranted in making my assertion in the first place.

My aim in this paper is not to evaluate the many proposals that have been advanced to account for such data (although there will be occasion to discuss some of these proposals in §§3.3–3.4). My claim is simply the following: we should strive for a unified explanation of the eavesdropping data about reason claims and modal claims. In both cases, we need to explain how an eavesdropper is in a position to reject an assertion the speaker was nevertheless warranted in making. Theorists have used such data to motivate a variety of proposals about the distinctive semantics of epistemic modals. The parallel data about reason claims suggests these expressions will have some kind of structurally similar semantics.

3.2.2 Embedding Data

Suppose Joe lives in a town where a murder has just occurred. The police have begun their investigation, but so far they have no suspects. Compare:

(3) # Joe is the murderer, but he might not be the murderer.[8]
(4) # Joe is the murderer, but there's no reason to believe he's the murderer.

Both sentences are extremely odd things to assert. However, these sentences differ sharply in their "embedding" behavior—i.e. they differ in how they contribute to complex sentences of which they are a part. As Seth Yalcin (2007; 2011) observes, sentences like (3)—which he calls "epistemic contradictions"—are also incoherent to *entertain* or merely *imagine* as true:

(5) # Suppose Joe is the murderer, but he might not be the murderer.
(6) # Imagine a scenario in which Joe is the murderer, but he might not be the murderer.

[7] See Stanley (2005) for one example of this semantics.
[8] I use "#" to mark sentences that are grammatical but infelicitous.

But sentences like (4)—which I'll call "reason contradictions"—embed just fine:

(7) Suppose Joe is the murderer, but there's no reason to believe he's the murderer.

(8) Imagine a scenario in which Joe is the murderer, but there's no reason to believe he's the murderer.[9]

Here is a different pair of sentences that illustrates the same point:

(9) # Suppose your gas gauge is broken, but it might not be broken.

(10) Suppose your gas gauge is broken, but there's no reason to believe it's broken.

The second sentence sounds clearly better than the first.

The effect persists even if we add a bit of context immediately before giving the sentences:

John lives in a town where there has been an outbreak of a virus. Often times those infected with the virus initially show no symptoms.

(11) # Imagine a scenario in which John is infected, but he might not be infected.

(12) Imagine a scenario in which John is infected, but there is no reason to believe he is infected.

[9]cf. Skorupski (2010, 42): "[W]e take it for granted that there can be propositions that are true even though there is no reason to believe that they are true." Let me also add a few clarifications. I'll refer to sentences that exhibit any of the following forms as "reason contradictions":

ϕ but there is no reason to believe that ϕ

ϕ but there is (every/good) reason to believe that not-ϕ

This terminology is not ideal since sentences exhibiting the first form do not have the form of contradictions whose second conjunct is within the scope of an operator, as in Yalcin's epistemic contradictions and our second type of reason contradictions. But it will be convenient to have a single term to cover both of the forms listed above, as both forms exhibit the same type of contrast with epistemic contradictions. Finally, I make no claim that embedded epistemic contradictions are *always* infelicitous. It may be possible to improve the felicity of these embeddings by adding certain contextual cues and reversing the order of the conjuncts (see Dorr and Hawthorne (2013), Dowell (ms.), Moss (2015), Sorensen (2009), Willer (2013), and Yanovich (2014)). Nevertheless, the fact that such maneuvers are needed in order to improve the felicity of the embeddings illustrates the contrast between sentences like (5)/(6) vs. (7)/(8): the latter require no such maneuvers in order to sound felicitous.

3.3 THE PUZZLE

We've seen that reason claims give rise to the very type of eavesdropping data that motivates a variety of proposals about the distinctive semantics of epistemic modals. We should thus expect that reason claims have some type of structurally similar semantics. However, the contrast in embedding data reveals that none of the leading semantics for modal claims can be extended to reason claims. Let me explain.

3.3.1 Two Semantics for Modal Claims

There are two main approaches to the semantics of epistemic modals. On the first, which I will call the **Orthodox Semantics**, modal claims express propositions concerning how things stand with some contextually relevant epistemic state. We've already seen one version of this theory above: *it might be that* ϕ is true at a context just in case what the speaker knows is compatible with ϕ. On this semantics, if I say "it might rain tomorrow", the sentence expresses the proposition that everything I know leaves open the possibility of rain. The **Orthodox Semantics** comes in many other varieties that differ as far as which *type* of epistemic state is at issue—knowledge, certain knowledge, or evidence—as well as *whose* epistemic state is at issue—the speaker's, some group's, or some larger epistemic state that includes the information the speaker or group is in a position to know.[10] But on all of these semantics, modal claims just describe what the world is like by describing how things stand with people's epistemic states: *it might be that* ϕ describes the world as being such that the contextually relevant epistemic state is compatible with ϕ.

The **Orthodox Semantics** contrasts with what I will call the **Heterodox Semantics** for modal claims. On the latter, modal claims do not express ordinary propositions at all—that is, they do not describe what the world is like. Instead, modal claims express first-order properties an information state might have ("information state" is just a general term for a state of belief, knowledge, or supposition; I'll explain what a "first-order" property is in a

[10] See Hawthorne (2012), Littlejohn (2011), and Dougherty and Rysiew (2009) for versions of the orthodoxy based on knowledge, certain knowledge, and evidence, respectively. See Stanley (2005), Dowell (2011), and DeRose (1991) for versions of the orthodoxy based on the speaker's knowledge, group knowledge, and accessible knowledge, respectively.

moment). Here is one version of this theory: *it might be that* ϕ expresses the property of an information state's being compatible with ϕ. If everything I believe is compatible with ϕ, then my belief state has this property—namely, being compatible with ϕ—and the modal claim is *true of* my belief state.

This type of semantics is typically used to develop relativist or expressivist theories of epistemic modal discourse.[11] For example, on a relativist theory like John MacFarlane's (2011; 2014), modal claims vary in their truth depending on the context from which they are assessed. If I say "it might rain", what I said will be true as assessed by me if what I know is compatible with rain, while the very same modal claim will be false as assessed by you if what you know rules out the possibility of rain. The modal claim is effectively information-neutral: instead of describing how things stand with some particular information state—the speaker's, a group's, etc.—the modal claim leaves open which information state matters and so is true relative to some information states and false relative to others. Finally, the truth of the modal claim at an information state just turns on how that information state itself bears on ϕ—say, by being compatible with ϕ, or by ruling out ϕ. This is the sense in which the property determined by the modal is "first-order".

Let's turn now to the question of how one might extend the **Orthodox Semantics** or the **Heterodox Semantics** to reason claims.

3.3.2 Problems for Extending the Heterodox Semantics

Defenders of the **Heterodox Semantics** will often motivate this theory by its ability to capture the type of eavesdropping data we discussed in §3.2.1. Suppose I don't know where the keys are and I say "The keys might be under the doormat". On MacFarlane's relativism, I am warranted in making this assertion because the modal claim is true relative to the context of assessment I presently occupy: what I know is compatible with the keys being under the doormat. However, if you've already checked the doormat, you are entitled to

[11] See, for instance, MacFarlane (2011, 2014), Moss (2013, 2015), Rothschild (2012), Swanson (2006, 2011, 2016), Willer (2013), and Yalcin (2007, 2011, 2012*a*); Egan's (2007) relativism employs what is essentially a notational variant of the **Heterodox Semantics**. But let me be clear: the **Heterodox Semantics** is not the only semantic framework for developing relativism. One of the morals of this paper is that we need to take seriously the possibility of a non-standard version of relativism. This will become clearer after §3.4.3; see Appendix A for the details.

reject my assertion because what I said is false at the context of assessment you presently occupy: what you know is incompatible with the keys being under the doormat.

A structurally parallel semantics allows us to explain the eavesdropping data about reason claims in exactly the same way. Suppose *there is* (*every*/*good*/*little*) *reason to believe that* ϕ determines a first-order property of an information state: that of counting in favor of believing that ϕ.[12] A reason claim will then be true of an information state s just in case s counts in favor of believing that ϕ.[13] There is no consensus yet on how exactly to understand this sort of counting-in-favor-of relation, but the following is a simple stand-in that will suffice for our purposes: s counts in favor of believing that ϕ just in case s makes it reasonably likely that ϕ.

Let's apply this semantics to the case discussed in §3.2.1. Recall that the DA learns that large quantities of drugs were found in Pete's car and on this basis claims there is good reason to believe that Pete is a drug trafficker. But Detective Barnes discovers that the drugs were planted and so rejects the DA's assertion. A relativist deployment of our above semantics explains why each party is warranted in saying what she did. The DA is warranted in making her assertion because the reason claim is true relative to her state of information: what she knows makes it reasonably likely that Pete is a trafficker. But Barnes is warranted in rejecting the DA's assertion because the reason claim is false relative to Barnes' state of information: what Barnes knows does not make it reasonably likely that Pete is a trafficker. The **Heterodox Semantics** for modal claims would thus appear to offer an attractive model for the semantics of reason claims.

However, there is a problem: any plausible, heterodox-style semantics for reason claims will inevitably make the wrong predictions about how these expressions embed.

[12]The differences between the grades of *every reason*, *good reason*, and *little reason* will plausibly reflect the degree to which the information state counts in favor of believing that ϕ. I will set aside these complications in what follows. Furthermore, it is plausible there exist uses of reason claims on which they characterize only a proper subset of a given information state as counting in favor of believing that ϕ. This reading is easier to access in sentences like the following: "I grant that there is *some* reason to believe that P, but in fact, I think that not-P." I will set aside this reading in what follows. But even if one maintains—implausibly—that reason claims *only* or *typically* characterize proper subsets of an information state, the problem I raise below for extending the **Heterodox Semantics** remains.

[13]Kolodny (ms.) endorses a similar semantics. Finlay (2014, Chapter 4) and Henning (2014) hold a similar view about relational reason claims. They take sentences of the form ϕ *is a reason to believe that* ψ to be true relative to some background information state s.

To understand this problem, it will be helpful to first see how the **Heterodox Semantics** explains the infelicity of embedded epistemic contradictions. Recall that epistemic contradictions have the following general form: *ϕ and might not-ϕ*. Recall also that on the **Heterodox Semantics**, *might not-ϕ* does not express a proposition about people's epistemic states. The sentence just determines a first-order property of an information state: that of being compatible with *not-ϕ*. This account of the semantic function of *might not-ϕ* pairs naturally with a view about the state of mind of believing, supposing, or, in general terms, accepting *might not-ϕ*: to accept this sentence is not to have an attitude towards a proposition *about* people's epistemic states—accepting the sentence just involves *being* in an information state that has the property expressed by the sentence. That is, to suppose that *might not-ϕ* is just to *be* in a suppositional state that leaves open the possibility of *not-ϕ*.[14] Now consider an embedded epistemic contradiction: *suppose that ϕ and might not-ϕ*. To suppose that ϕ is to be in a suppositional state that rules out the possibility of *not-ϕ*. To suppose that *might not-ϕ* is to be in a suppositional state that leaves open the possibility of *not-ϕ*. No coherent suppositional state has both of these properties—that of both ruling out and leaving open *not-ϕ*—which explains why embedded epistemic contradictions sound odd.[15]

Now consider the state of mind of supposing a reason contradiction. On our heterodox-style semantics from above, a reason claim determines the following property of an information state: that of counting in favor of believing that ϕ. Accordingly, to suppose there is reason to believe that ϕ is not to suppose some proposition *about* people's epistemic states—the supposition just involves *being* in a suppositional state whose content counts in favor of believing that ϕ.[16] But notice: if you suppose that ϕ, you are supposing something that

[14] Yalcin (2011) also requires that the attitude holder be sensitive to the question of whether *not-ϕ*. This qualification will not affect the discussion that follows.

[15] See Yalcin (2007) for the further details of this explanation.

[16] Greco (2014, 211) endorses a similar view about the mental state of believing one ought to believe that ϕ: "[T]he judgment that one ought to believe that ϕ is understood as a species of belief that ϕ". In other words, the belief that one ought to believe that ϕ does not take as its object some proposition about people's knowledge; the belief is just a special way of being related to ϕ itself.

counts conclusively in favor of believing that ϕ—namely, ϕ itself.[17] Thus, to suppose that ϕ is to suppose that there *is* reason to believe that ϕ. Hence, it is incoherent to suppose both that ϕ and that there is *no* reason to believe that ϕ. But of course, this is absurd. As we saw in §3.2.2, there is nothing incoherent about entertaining a scenario in which (say) one's gas gauge is broken but there is no reason to believe that the gauge is broken.

The problem generalizes to afflict any heterodox-style semantics for reason claims that satisfies the following two constraints:[18]

(a) If *there is reason to believe that* ϕ is true of an information state s, then *there is no reason to believe that* ϕ is false of s.

(b) If s entails that ϕ, then *there is reason to believe that* ϕ is true of s.

(a) follows from a basic point about the logic of epistemic reasons: there being reason to believe that ϕ is inconsistent with there being no reason to believe that ϕ. (b) is extremely plausible: surely an information state's entailing that ϕ supplies at least reason to believe that ϕ. However, (a) and (b) together ensure that it is incoherent to entertain a reason contradiction. To suppose that ϕ is to get into an information state s that rules out the possibility of *not-*ϕ and so entails that ϕ. Thus, by (b), *there is reason to believe that* ϕ will be true of s. Hence, by (a), *there is no reason to believe that* ϕ will be false of s. Whence, there is no coherent suppositional state s that both accepts that ϕ and that there is no reason to believe that ϕ. But this result is absurd. So reason claims do not have a heterodox-style semantics.

It is a surprising fact that we cannot extend the **Heterodox Semantics** to reason claims. MacFarlane and others had explicitly designed the **Heterodox Semantics** to account for the eavesdropping data about modal claims. But as we've just seen, this semantics is actually *inapplicable* to the case of reason claims, even though reason claims generate the very same type of eavesdropping data.

[17] One might object that ϕ is not a reason for believing that ϕ. But this claim runs counter to a variety of theories about the counting-in-favor-of relation and related views about evidential support. The conditional probability of ϕ given ϕ is 1, so as long as the prior probability of ϕ is less than 1, Bayesians will take ϕ to confirm ϕ. Similarly, ϕ entails that ϕ, so hypothetico-deductivists will take ϕ to confirm ϕ. And many epistemologists hold that in cases of veridical perception or veridical introspection, one's reason for believing that ϕ is simply ϕ itself (see, for instance, McDowell (1994) and Neta (2011)).

[18] See Appendix B for a formal proof of this point.

3.3.3 Problems for Extending the Orthodox Semantics

Here is where we are in the dialectic: we are looking for a structural parallel in the semantics of reason and modal claims capable of supporting a unified explanation of the shared eavesdropping data. And we've just seen that this structural parallel cannot be found by extending the **Heterodox Semantics** to reason claims.

Suppose we try instead to extend the **Orthodox Semantics** to reason claims. On this view, reason and modal claims would have a structurally parallel semantics in the following sense: both claims would express propositions concerning how things stand with some contextually relevant epistemic state. *It might be that* ϕ would describe the relevant state as being compatible with ϕ, while *there is reason to believe that* ϕ would describe the relevant state as counting in favor of believing that ϕ.

Now, if reason claims have this type of semantics, it is no mystery why reason contradictions are coherent to entertain: to suppose a reason contradiction would just be to entertain a scenario in which ϕ is true but someone's epistemic state does not count in favor of believing that ϕ.

However, epistemic contradictions describe extremely similar scenarios, if the **Orthodox Semantics** is correct: on this semantics, an epistemic contradiction just describes a scenario in which ϕ is true but someone fails to know that ϕ. Thus, if reason and modal claims both have an orthodox-style semantics, why are epistemic contradictions far more difficult to entertain than reason contradictions?

The problem runs deeper. Defenders of the **Orthodox Semantics** have devoted a great deal of attention to the task of explaining why embedded epistemic contradictions sound odd.[19] If "the gauge might not be broken" just means that someone fails to know it is broken, then it's not clear why there is an obvious difference in felicity between sentences like the following:

(13) # Imagine a scenario in which the gauge is broken, but it might not be broken.

(14) Imagine a scenario in which the gauge is broken, but Jake doesn't know it's broken.

[19] See Crabill (2013), Dorr and Hawthorne (2013), Ninan (forthcoming), Silk (2017), Sorensen (2009), Stojnić (2016, forthcoming), and Yanovich (2014).

The existing responses to this problem typically draw on some aspect of the pragmatics of epistemic modals. It is claimed that the contextual flexibility of these expressions accounts for their contrast with overt descriptions of knowledge, as in (14). However, if reason claims *also* describe contextually relevant epistemic states, then all of these pragmatic theories would seem to be false—for reason contradictions embed just fine.

Let me give one specific example of this problem. Some theorists claim that embedded epistemic contradictions sound bad because it's not clear whose epistemic state is at issue, as compared with sentences like (14).[20] But embedded reason contradictions also do not specify whose epistemic state is at issue, and yet they sound fine:[21]

(15) Imagine a scenario in which the gauge is broken, but there's no reason to believe it's broken.

The general problem is this: if we extend the **Orthodox Semantics** to reason claims, it's hard to see how any pragmatics could engage with modal claims but not reason claims so as to predict their difference in embedding behavior. After all, on this proposal, reason and modal claims would both be context-sensitive expressions that describe how things stand with some relevant epistemic state. Thus, while our extension of the **Orthodox Semantics** explains the embeddability of reason contradictions, the view yields an incoherent picture of the language of subjective uncertainty as a whole.

This is a surprising result. The **Orthodox Semantics** remains a popular theory, and as we noted above, many authors identify reasons for belief with items of knowledge. In fact, John Hawthorne and Jason Stanley each defend *both* of these theories.[22] But it turns out that we cannot integrate these views to provide a unified account of the semantics of reasons and modals.

[20] See Crabill (2013) and Dorr and Hawthorne (2013).

[21] Dorr and Hawthorne have a move available to them here. Their explanation of the epistemic contradiction data also relies on modal claims' having what they call "constrained interpretations". Dorr and Hawthorne might therefore deny that reason claims have such interpretations, which could in turn account for the contrast in embedding data. However, this only deepens the puzzle, for Dorr and Hawthorne also rely on these constrained interpretations to explain eavesdropping data (see Dorr and Hawthorne (2013, 880) and Hawthorne (2007)).

[22] See Hawthorne (2012), Hawthorne and Stanley (2008), and Stanley (2005).

We arrive now at our puzzle: What should the semantics of reason claims be? These expressions give rise to the distinctive pattern of eavesdropping data that has led numerous theorists to propose novel accounts of the semantics of epistemic modals. All of this suggests that reason claims have some kind of structurally parallel semantics. However, the difference in the embedding data makes it hard to see what this parallel could possibly be: neither the **Orthodox Semantics** nor the **Heterodox Semantics** can supply a unified semantics for reason and modal claims.

3.4 A SOLUTION

Any solution to our puzzle will involve constraining the semantics of reasons and modals along two different dimensions. First, we must identify some kind of structural parallel capable of supporting a unified explanation of the shared eavesdropping data. Second, this structural parallel must be such as to allow for the differences in embedding data.

Here are the views I will defend:

Mentalistic Semantics for Epistemic Reasons: Reason claims describe how things stand with the relevant knowledge, or more broadly, the information that lies within a thinker's "epistemic reach" (Egan (2007)).

Veritic Semantics for Epistemic Modals: Modal claims describe how things stand with the relevant truths, where there is no requirement that these truths be known or otherwise related to anyone's epistemic state.

If these views are correct, then reason and modal claims have the following structural parallel: both describe how things stand with relevant bodies of information. I will argue that this structural parallel is what accounts for the similarity in eavesdropping data. Nevertheless, each expression describes different *types* of information—knowledge vs. mere truths—and this difference is what accounts for the difference in embedding data.[23]

[23]Strictly speaking, **Mentalistic Semantics** is consistent with a variety of views about exactly what type of epistemic state reason claims describe—e.g. evidence, knowledge, certain knowledge, or accessible knowledge. But for the sake convenience I will often gloss this theory as holding that reason claims describe knowledge.

3.4.1 Characterizing Each Semantics

Mentalistic Semantics should seem like a familiar theory. All we have done here is transpose the central idea behind the **Orthodox Semantics** for modal claims to the case of reason claims: the latter describe relevant epistemic states.[24]

There is a straightforward motivation for this semantics that was already implicit in our discussion of reason contradictions above. Notice that there are striking parallels between reason contradictions like (16) and Moore-paradoxical sentences like (17):

(16) # The gauge is broken, but there's no reason to believe it's broken.
(17) # The gauge is broken, but I don't know it's broken.

Both sentences sound bad when asserted, but they describe possibilities it is coherent to entertain:

(18) Suppose the gauge is broken, but there's no reason to believe it's broken.
(19) Suppose the gauge is broken, but I don't know it's broken.

A simple explanation of these parallels is that reason claims leave implicit what Moore-paradoxical sentences make explicit: both describe how things stand with some relevant body of knowledge. Each is coherent to entertain because each describes failures of knowledge. Each is a bad thing to assert because asserting that ϕ represents oneself as knowing that ϕ, and one's knowing that ϕ contradicts one's not knowing that ϕ or there being *no* reason to believe that ϕ.

Now, we've already seen the tension in holding that reason and modal claims both describe knowledge: if both have this function, then it's extremely difficult to see how one could explain their difference in embedding behavior. However, there is an option we haven't yet considered: suppose one adopts an orthodox-style semantics for reason claims and the **Heterodox Semantics** for modal claims—i.e. reason claims are context-sensitive descriptions of knowledge, while modal claims determine first-order properties of information states. Call this the **Split View**.

[24] One caveat: I want to leave open the possibility that reason claims characterize epistemic states without literally describing them by expressing propositions *about* particular epistemic states. The possibility of this type of theory will only become clear in Appendix A, so I will set it aside for now.

The **Split View** is a step in the right direction. As we discussed in §3.3.2, the **Heterodox Semantics** accounts for the infelicity of embedded epistemic contradictions, and as we just discussed above, a knowledge-describing semantics for reason claims accounts for the felicity of embedded reason contradictions. Thus, the **Split View** captures the difference in embedding data.

However, by sacrificing a structural parallel in the semantics of reason and modal claims, the **Split View** fails to provide a unified explanation of the eavesdropping data. If the **Split View** were correct, we would need to come up with two explanations of the eavesdropping data: one along contextualist lines to account for the data about reason claims, and one along relativist or expressivist lines to account for the data about modal claims. But the existing models for each type of explanation work in fundamentally different ways.[25] It would be very unusual if the same pattern of data had such different explanations. Perhaps we could live with this result if this were the best we could do. But it isn't.

Suppose instead that modal claims have a *veritic* semantics: they describe how things stand with a set of relevant *truths*, where there is no requirement that these truths be known or otherwise related to anyone's epistemic state.[26]

Let me introduce this semantics by way of a comparison with the **Orthodox Semantics**. Take a simple example:

(20) It might rain tonight.

According to the **Orthodox Semantics**, a use of (20) expresses a proposition whose truth at a world w turns on how things stand with the contextually relevant individual or group's epistemic state at w.[27] Such a proposition divides up the space of possible worlds according to how things stand with people's knowledge at those worlds: a world where the relevant body of knowledge is compatible with rain is a world ruled in by the proposition, whereas a world where the relevant body of knowledge is inconsistent with rain is a world ruled out by the proposition. This is what it means for the modal claim to "describe" a body of knowledge.

[25] See, for instance, the dispute between von Fintel and Gillies (2008, 2011) and MacFarlane (2011, 2014).
[26] See Marushak and Shaw (ms.) for development and defense of this view.
[27] For the sake of simplicity, I will henceforth take the relevant epistemic state to be a state of knowledge.

Veritic Semantics treats (20) differently. On a simple, contextualist version of this semantics, a use of (20) expresses a proposition whose truth at a world w turns on how things stand with the contextually relevant true propositions at w. Examples of such propositions might be: the true propositions describing the shape of the clouds overhead and the rates of approaching storm fronts (say).[28] The proposition expressed by (20) will then be true at a world w just in case the shape of the clouds and rates of approaching storm fronts at w are consistent with rain. Such a proposition divides up the space of possible worlds according to how things stand with the meteorological facts at those worlds: a world where the relevant facts are compatible with rain is a world ruled in by the proposition, whereas a world where the relevant facts are inconsistent with rain is a world ruled out by the proposition.

Veritic Semantics and the **Orthodox Semantics** thus give two very different accounts of the proposition expressed by (20). On the first, the truth of this proposition at a world simply turns on how things stand with the meteorological facts at this world. On the second, the truth of the proposition at a world turns on how things stand with what people know at this world. Plausibly, there are possible worlds where the relevant knowers fail to know all of the relevant truths, so these two accounts of the proposition expressed by (20) are not equivalent.

Still, let me caution against one potential misunderstanding of **Veritic Semantics**. It might seem that defenders of this view must deny that knowledge plays a role in the interpretation of epistemic modals. But this is a mistake. It is entirely consistent with **Veritic Semantics** that a modal claim describes a particular set of truths *precisely because* these propositions are the ones known by the speaker to be true at the world of the context of use. For example, suppose someone is staring at the clouds and says "it might rain tonight". **Veritic Semantics** gives a natural characterization of the proposition expressed by this utterance: the speaker is proposing that the shape of the clouds is consistent with rain. This proposition will be true at a world depending on how things stand with the shape of the

[28]Let me temporarily set aside the question of how the context of use determines which propositions are relevant. This is a challenging question about the pragmatics of language that arises also for defenders of the **Orthodox Semantics**—after all, they must explain how the context of use settles which body of knowledge is relevant. But just to fix ideas, one might hold that the relevant truths are determined in part by the speaker's intentions, the aims of the conversation, and the propositions salient in the context as being true. I'll have more to say about contextual relevance below in §§3.4.2–3.4.3.

clouds at that world. But notice: it is highly plausible that the speaker's utterance expresses this proposition because she is aware of and attending to the shape of the clouds. The key point here is that the speaker's knowledge can influence which proposition is expressed by a modal claim without the modal claim literally describing the speaker's knowledge. There will be many worlds at which the speaker is entirely ignorant of the shape of the clouds, but the truth of the modal claim at these worlds still just turns on how things stand with the shape of the clouds themselves. This is what it means to say that the proposition describes the meteorological facts, not the speaker's knowledge.

Now, some might object that we cannot make sense of how a speaker is warranted in asserting a modal claim unless this claim describes the speaker's knowledge. But consider an ordinary factual claim like "it's raining". The proposition expressed by this sentence does not describe the speaker's knowledge. Rather, the proposition is true at a world just in case it is raining at this world, and there will be many such worlds at which the speaker is entirely ignorant of whether it's raining. But of course, it does not follow that no one is ever warranted in asserting that it's raining. All it takes for one to be warranting in asserting "it's raining" is that one know that it is raining *at the world in which one uses this sentence*. Similarly, all it takes for one to be warranted in asserting "it might be raining" is that one know that the relevant meteorological facts *at the world in which one uses this sentence* are compatible with rain. Thus, the possibility of warranted assertion does not require modal claims to literally describe the speaker's knowledge.

However, it is even consistent with **Veritic Semantics** that a modal claim does occasionally describe the speaker's knowledge. The interests of the speaker or the larger aim of the conversation may well support a reading of a modal claim on which the relevant truths are those concerning what the speaker knows. The central claim of **Veritic Semantics** is just that there is no *default* preference for describing mentalistic truths of this sort. Truths about what people know are a small subset of the many truths one might be interested in characterizing by means of a modal claim. So we should expect such mentalistic truths to show up only occasionally as the relevant truths that modal claims describe.

To sum up: On the view I defend, it is reason claims, not modal claims, that characteristically function to describe bodies of knowledge. The difference between each semantics lies

in the proposition expressed by a typical use of each sentence: the truth of a reason claim turns on how things stand with what is known at a world, while the truth of a modal claim turns on how things stand with the facts at a world—facts which may or may not be known by any contextually relevant individuals.

Let's now consider how this proposal resolves the puzzle of the previous section.

3.4.2 Revisiting the Embedding Data

According to **Veritic Semantics**, modal claims describe how things stand with the relevant truths. But relevance is often a matter of salience. Suppose I am dining with Mary and her friends, and we are asked if anyone wants coffee. I reply, "No one wants coffee." It is clear I have thereby conveyed that Mary does not want coffee. On the standard account of this phenomenon, the phrase "no one" quantifies over a set of relevant persons, and Mary's salience in the context makes her one of these persons.[29] Thus, if none of these persons wants coffee, then neither does Mary.

This theory also explains why the following sentence sounds contradictory:

(21) # Mary wants coffee, but no one wants coffee.

The first conjunct makes Mary a salient person, so she figures in the domain of the subsequent quantifier "no one"—i.e. the first conjunct "restricts" the domain of the quantifier in the second. Thus, any world at which the first conjunct is true will be a world at which the second is false.

Veritic Semantics makes an analogous prediction about epistemic contradictions like (22):

(22) # The gauge is broken, but it might not be broken.

Modals are standardly understood as quantifiers whose domains are restricted by sets of propositions.[30] The distinctive claim of **Veritic Semantics** is that modal domains are restricted by relevant truths—these truths need not be known or otherwise related to anyone's epistemic state. Now return to (22): to assert that the gauge is broken is to make this

[29] See, for instance, von Fintel (1994) and Stanley and Szabó (2000).
[30] See Kratzer (1977, 1981, 2012).

proposition salient as being true. But again, relevance is often a matter of salience. So we should expect the proposition that the gauge is broken to restrict the domain of the subsequent modal—after all, the modal is characterizing relevant truths and the proposition in question is salient as being true. Hence, asserting (22) describes the world as being such that the gauge is broken and that it is compatible with the relevant truths—including the fact that the gauge is broken—that the gauge is not broken. But no world has this property, which is why the sentence sounds odd.

Notice that such domain restriction effects persist in suppositional environments:

(23) # Suppose Mary wants coffee, but no one wants coffee.

Here again Mary figures in the domain of "no one". So it is no surprise that embedded epistemic contradictions continue to sound contradictory:

(24) # Suppose the gauge is broken, but it might not be broken.

(24) asks us to entertain a scenario in which it is true that the gauge is broken and in which it is compatible with the relevant truths—including the fact that the gauge is broken—that the gauge is not broken. But there are no such scenarios, so the sentence sounds odd.[31]

Contrast the situation with reason contradictions:

(25) Suppose the gauge is broken, but there's no reason to believe it's broken.

If reason claims describe relevant knowledge, then supposing that the gauge is broken should not suffice to make this information part of the information that matters for assessing the truth of a subsequent reason claim. For to merely suppose that the gauge is broken is not to suppose that anyone *knows* or *has evidence* for thinking that the gauge is broken. It is therefore easy to access a felicitous reading of (25) on which the sentence just asks us to entertain a scenario in which the gauge is broken but in which the epistemic state of some individual or group fails to count in favor of believing that the gauge is broken. Thus, it is the difference in the *type* of information described by reason and modal claims that explains the difference in embedding behavior.

[31] See Marushak and Shaw (ms.) for discussion of additional parallels between epistemic contradictions and what we call *quantified contradictions*, i.e. sentences like "Mary wants coffee, but no one wants coffee".

The key to this explanation lies in the fact that we've identified different default constraints on the interpretation of reason and modal claims. There is no general mechanism by which the information that the gauge is broken restricts the domain of any information-describing expression that comes later in the sentence. Were there such a mechanism, embedded reason contradictions would sound bad as well.[32] Instead, the information that the gauge is broken restricts the domain of subsequent epistemic modals only because these modals by default characterize mere truths. In contrast, reason claims by default characterize relevant epistemic states, so the mere salience of the gauge's being broken does not suffice for this information to count as relevant for the truth of the reason claim: the reason claim is characterizing relevant epistemic states, but the gauge can be broken without anyone being in a position to know this fact. So again, it is the difference in the type of information each expression describes that accounts for the difference in embedding behavior.

3.4.3 Revisiting the Eavesdropping Data

While reason and modal claims differ in which type of information they characterize, there is nevertheless an underlying unity in their basic semantic function: both describe how things stand with relevant bodies of information. I propose that this unity in basic function is what accounts for the common pattern in eavesdropping data.

Notice that many of the existing explanations of the eavesdropping data about modals make no essential reference to which *type* of information these expressions describe. For example, on the "cloudy contextualism" of von Fintel and Gillies (2011), a use of a modal claim "puts into play" a set of propositions describing how things stand with different bodies of knowledge.[33] One proposition describes how things stand with the speaker's knowledge, while another describes how things stand with a larger group's knowledge. The speaker is warranted in asserting a modal claim as long as she is warranted in asserting at least one of these propositions. But a listener is warranted in rejecting a modal claim as long as she

[32] The behavior of reason contradictions thus falsifies the accounts of epistemic contradictions given in Silk (2017) and Stojnić (2016, forthcoming), as these accounts rely precisely on such general purpose domain restriction mechanisms.

[33] More precisely, it is uses of "bare" modals—i.e. modals whose domains are not already restricted by some further element of the discourse—that put into play sets of propositions.

is warranted in denying the most informative proposition about which she has an opinion. This theory explains why a listener with more information can reject an assertion that the speaker was nevertheless warranted in making. For suppose that the speaker is warranted in asserting one of the weaker propositions in the set, such as the one that concerns her own body of knowledge. And suppose the listener is warranted in denying one of the stronger propositions in the set, such as the one that concerns how things stand with the group's body of knowledge. On the cloudy contextualist model, both parties are within their linguistic rights when the speaker makes an assertion that the listener rejects.

This explanation goes through just as well if modal claims describe relevant truths. A use of a modal claim will put into play a set of propositions whose members describe how things stand with different batches of truths. A speaker might be warranted in asserting one of the weaker propositions in this set, while a listener might be warranted in denying one of the stronger propositions in this set. And of course, a parallel story can be told for reason claims, as they have essentially the same semantics von Fintel and Gillies (2011) attribute to modal claims.

We thus arrive at a unified explanation of the eavesdropping data. If reason and modal claims both describe relevant bodies of information, it is open for us to hold that neither describes a *unique* body of information in a typical context of use. Both expressions might put into play sets of propositions, which in turn give rise to the characteristic pattern of eavesdropping data.

Let me emphasize that cloudy contextualism is just one option for explaining the eavesdropping data. There are others.[34] The important point here is that the semantics I've given allows for a unified explanation of the data and provides a clue as to the general shape it should take—namely, the explanation should rely on some common feature of information-describing expressions, one that is shared among expressions that nevertheless differ in which type of information they describe.

[34]See Appendix A for a non-standard relativist explanation of the data.

3.5 CONCLUSION: THE CONCEPT OF AN EPISTEMIC REASON

We began with a puzzle about the semantics of epistemic reasons. The shared eavesdropping data suggests that reason and modal claims have some type of structurally parallel semantics. But the difference in embedding data rules out extending each of the leading semantics for modal claims to reason claims. How, then, can we give a semantics for reason claims that accounts for both their similarities and differences with modal claims?

I argued that the solution to this puzzle lies in embracing a novel semantics for epistemic modals: it is reason claims that describe knowledge, while modal claims describe truths. This theory immediately explains the difference in embedding data. The theory also allows us to give a unified explanation of the similarity in eavesdropping data. And all of this takes place within a single semantic framework on which reason and modals claims both describe relevant bodies of information. None of the other accounts of epistemic modals can be integrated with the semantics of reasons to yield a theory with all of these virtues. So it turns out that the language of epistemic reasons has much to teach us about epistemic modals.

But do we learn anything about epistemic reasons as well? I will conclude with some brief reflections on this question. I will argue that the semantics of epistemic reasons ultimately helps us understand our *concept* of an epistemic reason.

Let me first make a point about epistemic modals. In hindsight, it should come as no surprise that modal claims describe relevant truths. When we engage in deliberation about whether it might be that ϕ, we typically do so because we are interested in the question of whether ϕ. The bearing of a set of true propositions is extremely relevant to this latter question, for if these propositions are inconsistent with ϕ, then ϕ must be false.

Furthermore, it should come as no surprise that modal claims *merely* describe relevant truths. There is no mentalistic restriction on quantifier domains generally. If I say "All of the students have left the classroom", it is easy to hear this sentence as false if a student is hiding in the closet.[35] What I said does not mean that all of the students *I'm aware of* left the classroom. Now compare: modals are just quantifiers of a special sort—namely,

[35] cf. Ichikawa (2011, 384).

quantifiers whose domains are restricted by sets of propositions. So it would be surprising if there were some mentalistic restriction on modal domains, whereby these domains are restricted only by the truths of which some thinker is aware. And indeed, *might* ϕ does not mean that all of the truths *I'm aware of* are compatible with ϕ. To be sure, one will only be warranted in asserting a claim about how things stand with the truths of which one is aware. But the mistake behind the **Orthodox Semantics** was to bake this fact about pragmatics into the semantics of epistemic modals. Once we consider embedded uses of modal claims, we see that these expressions characterize relevant truths regardless of whether these truths are connected to anyone's epistemic state. This is why one cannot easily entertain a scenario in which ϕ is true and it is possible that *not-*ϕ.

But then one wonders: Why don't reason claims merely describe relevant truths as well? Why is there a default preference built into the semantics of reason claims for describing bodies of knowledge, or information connected to thinkers' epistemic states?

My answer is that the semantics of epistemic reasons reflects of our concept of an epistemic reason. Philosophers will often distinguish between so-called objective reasons vs. the reasons one "has" or "possesses".[36] The first are simply true propositions, while the second are propositions to which a thinker enjoys some kind of epistemic or cognitive access—e.g. they are the propositions that one knows or perhaps merely believes. However, what I've essentially argued above is that there is a semantic default for speaking of possessed reasons when it comes to talk of reasons for belief in natural language: reason claims describe how things stand with the information connected to thinkers' epistemic states.[37]

I will now argue that this semantic default is conceptual in origin: our pre-theoretical concept of an epistemic reason is that of a piece of information that is known, easily known, believed, or otherwise related to a thinker's epistemic state. In short: our pre-theoretical concept of an epistemic reason is that of a possessed epistemic reason. Call this the **Conceptual Priority Thesis**.

[36]See, for instance, Comesaña and McGrath (2014), Henning (2014), Hornsby (2008), Lord (2010, 2015), Neta (2008*a*, 2011), Schroeder (2008, 2011), and Skorupski (2010, 108).

[37]Let me emphasize that I am using the term "epistemic state" broadly, so that it includes not only states of knowledge but also states of evidence, or even experiences. I use the term in this manner so as not to beg any questions about whether possessed epistemic reasons must be factive. See Comesaña and McGrath (2014), Lord (2010), and Schroeder (2008) for discussion.

The **Conceptual Priority Thesis** is a claim about the concepts we bring to philosophical inquiry. It is a claim about what we ordinarily take epistemic reasons to be. Of course, our pre-theoretical concept does not settle the final shape of a philosophical analysis of this concept. But neither can our pre-theoretical concept be ignored. Faithfulness to our pre-theoretical concept is one of the main criteria for evaluating a philosophical analysis of a concept (e.g. consider the role of Gettier cases in undermining the JTB analysis of knowledge). Our pre-theoretical concept also governs the metaphysical analysis of the property picked out by this concept (e.g. it would be a serious mark against an account of the nature of water if water turns out not to be the stuff that fills the oceans). So while the **Conceptual Priority Thesis** does not settle the metaphysics or conceptual analysis of reasons, this thesis will place important constraints on how we approach these issues.

Let me turn, then, to my argument for the **Conceptual Priority Thesis**. The argument is simple: this thesis offers the best explanation of the semantic default for speaking of possessed epistemic reasons. Consider some of the alternatives.

One explanation might appeal to the constraints on what it takes for someone to be warranted in asserting a reason claim. One is not warranted in making a claim about how things stand with reasons one doesn't possess. So perhaps there is a semantic default for describing the reasons one *does* possess. But this constraint is far too strong: we've already seen from the eavesdropping data in §3.2.1 that reason claims do not simply describe the speaker's knowledge—if they did, it's not clear how an eavesdropper with more information could ever be warranted in rejecting a speaker's assertion of a reason claim. The semantic default we seek to explain is not that reason claims describe the speaker's reasons but that reason claims describe *someone's* reasons—i.e. reason claims by default describe how things stand with information connected to thinkers' epistemic states.

One might argue instead that this semantic default is merely an accident of the English language. It is a complex empirical question whether this is correct. But let me offer some reasons for doubt. Talk of reasons for belief is not unique to English. And indeed, native speakers of Cantonese, Dutch, German, Italian, and Spanish have confirmed that these languages contain roughly synonymous expressions. So at present, an appeal to linguistic accident is mere speculation.

Another explanation might appeal to a claim about concept acquisition: perhaps the semantic default is no accident because we acquire the concept of an objective reason only after we acquire the concept of a possessed reason. But the problem here is that the order of concept acquisition frequently fails generate semantic defaults for interpretations involving the initially acquired concept. For example, there is a consensus in developmental psychology that children acquire the concept of epistemic modality after grasping the concept of deontic modality (i.e. modality concerning obligations and permissions).[38] But modals like "might" have a default epistemic interpretation, while other modals have no default interpretation one way or the other.

Compare now the explanation offered by the **Conceptual Priority Thesis**: talk of epistemic reasons is talk of knowledge, evidence, or information within a thinker's epistemic reach because this is what we ordinarily take epistemic reasons to *be*. This theory is simple, intuitive, and explains why it is no accident that a language would include such a default.

However, the **Conceptual Priority Thesis** is not entirely innocent. Indeed, the thesis calls into question widely held views about the nature of reasons. For example, according to T.M. Scanlon, reasons in general are just facts that count in favor of things, so reasons for belief are just facts that count in favor of belief.[39] Consider also what Mark Schroeder calls the **Factoring Account**. On this account, possessed reasons are objective reasons that we possess—i.e. the concept of a possessed reason is factored into the concepts of an objective reason and the possession relation.[40]

The **Conceptual Priority Thesis** suggests that both of these accounts are getting things backwards. On our pre-theoretical understanding of reasons for belief, it is not as though the world contains reasons for belief that are out there, hidden, and awaiting our discovery. Rather, facts are only intelligible as reasons for belief in relation to a thinker who is aware of or could easily come to be aware of these facts. It is therefore a mistake to analyze the reasons we possess as facts that are reasons which we in turn possess.

[38] See Papafragou (2002) for an overview of this literature.

[39] See Scanlon (1998, 17) and Scanlon (2014, 30), as well as Alvarez (2010, Chapter 2), Finlay (2014, 85), Parfit (2011, 31–32), and Reisner and Steglich-Petersen (2011, 1).

[40] See Schroeder (2008, 2011). Note that Schroeder himself rejects the **Factoring Account**. Lord (2010) defends the account against Schroeder's objections. See Schroeder (2008) for discussion of the role the account may have implicitly played in the development of 20$^{\text{th}}$-century epistemology.

The concept of an objective epistemic reason is plausibly understood as a philosophical abstraction based on the more fundamental notion of a possessed epistemic reason. We start by noticing that some of our beliefs are based on reasons that count in favor of holding these beliefs. We then isolate the counting-in-favor-of-belief relation and ask what it takes for this relation to obtain: some take the relation to be probabilistic, others take it to be explanatory, and so on. We finally introduce the concept of an objective epistemic reason to mark off the true propositions that stand in the counting-in-favor-of-belief relation.

However, the **Conceptual Priority Thesis** should caution us from simply classifying these propositions as reasons for belief themselves. On our pre-theoretical concept of an epistemic reason, facts only count as reasons for belief if they lie within the epistemic reach of a thinker.

The upshot is this: philosophers have widely misunderstood the relation between epistemic reasons and epistemic modals. They have taken the former to merely concern facts and the latter to concern knowledge. But in reality, it is the reverse that is true.

4.0 HOW TO FORMULATE FALLIBILISM

4.1 INTRODUCTION

Fallibilism commands near-universal acceptance in recent epistemology. But there is a surprising lack of consensus on how to formulate the view. For example, Cohen (1988) and Lewis (1996) both develop contextualist epistemologies on which *S knows that* ϕ is sometimes true even if S's evidence fails to entail that ϕ. But Cohen thinks this position vindicates fallibilism, while Lewis thinks fallibilism "mad". Consider also the recent debate over the semantics of so-called concessive knowledge attributions, or CKAs—i.e. sentences of the form *S knows that* ϕ*, but it's possible/it might be that* ψ, where ψ entails not-ϕ, and the modal in the second conjunct receives the epistemic interpretation.[1] Stanley (2005) denies that CKAs state the fallibilist thesis, while Reed (2013, 53) takes certain CKAs to be "obvious" consequences of the view. Leite (2010) concurs with Stanley, while Worsnip (2015, 225 n. 1) describes Stanley as attempting to "reconstrue fallibilism less ambitiously" so as to avoid the challenge posed by CKAs.

In this paper, I aim to clarify the formulation of fallibilism by defending the following general claim: fallibilism is just a thesis about the type of epistemic position consistent with knowledge, where a subject's "epistemic position" consists of the grounds upon which she might count as knowing that a proposition is true.[2]

[1] For now we can pick out the epistemic interpretation by ostension: it is the natural reading of *might* and *possible* in sentences like *The spare key might be under the doormat* and *It's possible that the teachers will strike*. The term "concessive knowledge attribution" is due to Rysiew (2001).

[2] I intend the "grounds of knowledge" to be interpreted broadly, so as to include both internalist and externalist types of justifications. However, the *grounds* of knowing that ϕ will exclude the knowledge that ϕ itself.

Some may think this formulation obvious and hardly worthy of extended defense. But the defense is urgently needed. Many epistemologists characterize fallibilism in other terms—e.g. as the view that knowledge is consistent with the possibility of error, or that one can know things that might be false, or that CKAs are sometimes true.[3] Notably, each of these formulations makes use of natural language epistemic possibility modals. It is widely held that epistemic modals describe how things stand with some individual or group's knowledge or epistemic position.[4] So the latter formulations might seem to be equivalent to the one I intend to defend. However, I will argue that there is no such equivalence: epistemic modals have no default function where they describe how things stand with anyone's evidence, knowledge, certain knowledge, or other epistemic state.

This point about epistemic modals has far-reaching consequences for how to theorize about fallibilism. Once we see that "epistemic" modals do not describe epistemic states, it becomes plain that epistemic modals have no place in the formulation of fallibilism. At best, they render the view epistemologically insignificant; at worst, they leave the view unrecognizable. A similar problem befalls several prominent arguments both for and against fallibilism. These arguments rely on intuitions about epistemic modals, so the arguments either fail to hit their mark, or they succeed by targeting only an epistemologically insignificant theory. In short: if we want to know whether fallibilism is true, we need a new methodology.

4.2 WHAT IS FALLIBILISM?

"Fallibilism" is a theoretical term, but its definition cannot be settled by stipulation. Any formulation of the view is answerable to a kind of rough and intuitive sense of the term, on which knowledge is compatible with human fallibility, broadly construed. Here is how DeRose (2017, 145) puts the point:

There seems to be an intuitive, but difficult-to-get-precise-about, sense in which we

[3]See, for instance, Dougherty (2011), Worsnip (2015), and Reed (2013), respectively.

[4]See, for instance, Anderson (2014), DeRose (1991, 1998), Dodd (2011), Dorr and Hawthorne (2013), Dougherty and Rysiew (2009, 2011), Dowell (2011), Fantl and McGrath (2009), Hacking (1967), Hawthorne (2004, 2012), Littlejohn (2011), Moore (1962), Stanley (2005), and Yanovich (2014).

humans are fallible with respect to everything, or at least nearly everything, that we believe, and "fallibilism" is sometimes used to designate this fact about us. But we will here be interested in uses of "fallibilism" in which it instead asserts that we can *know* things with respect to which we are fallible. "Intuitive fallibilism" can then be the position that knowing some fact is compatible with being fallible with respect to that fact in the murky-but-intuitive sense in question.

Intuitive fallibilism is our subject. Definitions of this term are best understood as what Quine, following Carnap (1947, §2), calls "explications": "We fix on the particular functions of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms to our liking, that fills those functions" (Quine (1960, 258–259)). That is, the definition we seek is not an attempt to spell out what Peirce had in mind upon introducing the term "fallibilism" into the lexicon.[5] Nor do we seek to characterize how philosophers are presently using this term. Rather, our aim is to provide a precise characterization of "fallibilism" that captures the aspects of our intuitive conception that epistemologists have found well worth troubling about for at least the past century, and arguably much longer.

I begin by introducing what are now the two standard ways of glossing fallibilism:

The Evidence Gloss: Fallibilism is a thesis about the type of epistemic position consistent with knowledge, e.g. that knowledge can be based on non-entailing evidence.

The Modal Gloss: Fallibilism is a thesis about the truth-conditions of sentences involving typical uses of natural language epistemic modals, e.g. that CKAs are sometimes true.

These glosses are not themselves definitions of fallibilism—they *figure* in definitions of fallibilism. More precisely: we can take each gloss to determine a property of theories. We can then define fallibilism by using predicates that express these properties.

For example, one version of the **Evidence Gloss** determines the following property: that of being a theory on which knowledge does not require entailing evidence. Let F denote this property. We can then use this predicate to define fallibilism as follows:

[5]See the OED entry on "fallibilism" for examples of Peirce's usage.

x is a fallibilist theory $=_{df}$ x is F.[6]

One might instead use both of our glosses to define fallibilism. Let G denote the property of being a theory on which CKAs are sometimes true. Then define fallibilism as follows:

x is a fallibilist theory $=_{df}$ x is F and x is G.

I'll discuss all of these options in more detail below. But let me first say a bit more about the glosses themselves.

The **Evidence Gloss** should be familiar. On this way of characterizing fallibilism, the view is at least in part a thesis about the strength of epistemic position required for knowledge (the term "evidence" in the name of the gloss is simply shorthand for "epistemic position"). For example, Stanley (2005) describes fallibilism as the view that knowledge can be based on non-entailing evidence, while Reed (2002) describes the view as one on which knowledge can be based on a justification that could have failed to yield knowledge.[7] In general terms, the **Evidence Gloss** characterizes fallibilism as a rejection of the stringent demands on knowledge typically employed in skeptical arguments.

The **Modal Gloss** is perhaps less familiar. On this account, fallibilism is at least in part a thesis about the truth-conditions of natural language epistemic modals. For example, Worsnip (2015) describes his aim as to "vindicate a form of fallibilism that is very robust indeed" and he clarifies his view as follows: "S can sometimes truly assert 'it is possible that not-p' even though S knows that p" (Worsnip (2015, 225–226)). Similarly, Reed (2013) claims that fallibilists are committed to the semantic consistency of CKAs like *S knows that ϕ, but it might be that not-ϕ*. As he puts it: "We cannot take these kinds of CKAs to be false in every instance without thereby rejecting fallibilism" (Reed (2013, 53).) One can also employ the **Modal Gloss** by *using*, not mentioning epistemic modals. Consider Dougherty's (2011) definition of fallible knowledge:

S fallibly knows that p $=_{df}$ (i) S knows that p, but (ii) p might (epistemically) be false.

[6]Strictly speaking, this is a definition of the term "fallibilist theory". But "fallibilism" is simply the doctrine that some fallibilist theory is true.

[7]Other authors employing versions of the **Evidence Gloss** include Cohen (1988), Dougherty (2011), Dougherty and Rysiew (2009), Fantl and McGrath (2009), Feldman (1981), Littlejohn (2011), Neta (2011), and Pryor (2000).

Trivially, if fallibilism is true, then there are metaphysically possible instances of fallible knowledge. But on Dougherty's definition, no one will ever have fallible knowledge if the truth of (i) precludes the truth of (ii).[8] Thus, if fallibilism is true, then epistemic modals need to have the right kind of truth-conditions. To generalize: Dougherty, Reed, and Worsnip are all proposing uses of the term "fallibilism" on which the view is at least in part a thesis about the truth-conditions of natural language modals.[9]

Let me clarify this last point. The **Modal Gloss** is a way of characterizing fallibilism by means of *natural language epistemic modals on their typical use*. The gloss is not an attempt to characterize fallibilism by means of terms like "epistemic possibility". The latter is a philosopher's term, not a piece of natural language. The **Modal Gloss** is also not an attempt to characterize fallibilism by means of some atypical use of epistemic modals present only in the mouth of the theorist or accessible only through some special contextual setup. When Dougherty, Reed, and Worsnip use modals to gloss fallibilism, they instruct the reader only to read the modals epistemically.

Still, one point of caution is in order. When assessing a proposed characterization of fallibilism, one must take particular care to distinguish between talk of what is possible and talk of what is possible *for some subject S*. The latter is ubiquitous in the literature on fallibilism, even in contexts where theorists are ostensibly giving a semantics for natural language possibility talk. Here is an example from Dougherty and Rysiew (2009, 127):

> q is epistemically possible for S iff not-q isn't entailed by S's evidence.

Strictly speaking, this biconditional characterizes something called "epistemic possibility for S". But it is clear from the context that Dougherty and Rysiew mean to characterize the truth-conditions of sentences like "it is possible that q", where the modal receives an epistemic reading. Their implicit assumption is this:

> "it is possible that q" (on the epistemic reading) is true iff q is epistemically possible for S.

[8] Presumably, Dougherty intends that (i) and (ii) be evaluated at the same context of use.

[9] See Anderson (2014), Dougherty and Rysiew (2009), Fantl and McGrath (2009), and Reed (2012) for other examples of the **Modal Gloss**.

As I'll argue below, this assumption is actually false. But for now, I simply want to draw attention to the fact that an assumption is being made in the first place. "possible for S" is functioning here as a theoretically-loaded phrase that reflects a particular view about the truth-conditions of epistemically modalized sentences—namely, that these sentences are true relative to how things stand with some subject S.

Consider, then, (1) and (2):

(1) Fallibilism is the view that knowledge is consistent with the possibility of error.

(2) Fallibilism is the view that S can know that ϕ even if it is possible for S that not-ϕ.

(1) is a version of the **Modal Gloss**, if we give "possibility" its natural epistemic reading. (2) is underspecified. If "possible for S" characterizes the state of S's epistemic position, then (2) counts as a version of the **Evidence Gloss**. But one might instead treat "possible for S" as a kind of shorthand for natural language talk of possibility.[10] In this case, (2) counts as a version of the **Modal Gloss**, although we must be careful not to read "for S" as having any theoretical significance.

Let me turn now to the question of how one might use our two glosses to define fallibilism. I will regiment such definitions as follows. Let "E-predicates" be predicates expressing properties determined by some version of the **Evidence Gloss**. Let "M-predicates" be predicates expressing properties determined by some version of the **Modal Gloss**. And define falllibilism by means of clauses with the following structure, where the right-hand side of the clause contains some E-predicate, M-predicate, or logical combination thereof:

x is a fallibilist theory $=_{df}$ ...[11]

Now, there are many E-predicates and M-predicates, and thus many possible definitions of fallibilism. My aim here is not to settle on any particular definition. Rather, I am con-

[10] See Worsnip (2015).

[11] I will assume that every term in a language receives at most one definition. So if one wishes to define fallibilism by means of more than one predicate, all must figure in the same definition clause. Note also: if every term receives at most one definition, then one cannot express the equivalence of two definitions by giving multiple definitions of the same term. Strictly speaking, a judgment of equivalence should be expressed as a thesis about the equivalence of two *languages*, each containing terms one takes to be defined by co-extensive, or co-intensive predicates (say).

cerned to assess the different views one might hold about the general shape such a definition should take. There are five views that will be the subject of my discussion.

The first is what I will call the **Equivalence Account**: there are E-predicates and M-predicates by which one can offer adequate and equivalent definitions of fallibilism. For example, suppose one holds the following semantics for epistemic modals:

> *it might be that* ϕ is true at a context-world pair $<c,w>$ iff the c-relevant S's evidence at w fails to entail that not-ϕ.

On this semantics, epistemic modals describe how things stand with the evidence of some contextually relevant subject. Now consider a CKA like "I know it's raining, but it might not be raining." A use of this sentence plausibly makes the speaker the contextually relevant subject. So by our above semantics, the second conjunct is true just in case the speaker's evidence fails to entail that it's raining. The first conjunct is true just in case the speaker knows it's raining. Thus, if the entire sentence is true, then there exists an instance of knowledge based on non-entailing evidence. In other words: any theory on which our CKA is sometimes true will be a theory on which knowledge does not require entailing evidence, and vice versa. Let us use our predicates F and G to denote these properties of theories. Anyone who endorses our above semantics will then view the following definitions as equivalent:[12]

> x is a fallibilist theory $=_{df}$ x is F.
> x is a fallibilist theory $=_{df}$ x is G.

In general terms, the **Equivalence Account** takes fallibilism to be a thesis about a knower's epistemic position and the truth-conditions of epistemic modals—but the two come to the same thing.[13]

Contrast the **Equivalence Account** with what I will call the **Conjunction Account**: fallibilism can only be properly defined by means of *complementary* E-predicates and M-predicates—i.e. predicates that are neither co-extensive nor mutually exclusive. For example, suppose one denies that epistemic modals describe the state of a subject's evidence. One

[12]More precisely: the definitions will be equivalent in the sense that they employ co-extensive predicates. Also, the definitions should be read as defining terms in two different languages. See n. 11.

[13]Defenders of the **Equivalence Account** include Dougherty (2011), Dougherty and Rysiew (2009), and Fantl and McGrath (2009).

might still hold that fallibilism is a theory on which CKAs are sometimes true. And one might also hold that fallibilism is a theory on which knowledge can be based on non-entailing evidence. If one takes these claims to offer an exhaustive characterization of the fallibilist thesis, then the resulting definition is consistent with the **Conjunction Account** but not the **Equivalence Account**.[14]

Consider next what I will call the **Ambiguity Account**: there is no single term, "fallibilism", that can be defined so as to capture the intuitions epistemologists are aiming to characterize. There is only fallibilism$_1$, fallibilism$_2$, and so on, where some of these terms are defined by E-predicates, others are defined by M-predicates, and the resulting definitions are not equivalent.[15]

Our final two accounts differ from all of the above in denying that E-predicates and M-predicates *both* figure in adequate definitions of fallibilism. According to what I will call **Evidence Purism**, fallibilism should defined by some E-predicate(s) and no M-predicate(s). On this account, fallibilism *just is* a thesis about the type of epistemic position consistent with knowledge. This is the view I will defend in this paper.[16]

One might instead hold that fallibilism should be defined by some M-predicate(s) and no E-predicate(s). I will call this view **Modal Purism**. On this account, fallibilism *just is* a thesis about the truth-conditions of sentences involving typical uses of natural language epistemic modals.[17]

To sum up, the two standard glosses on fallibilism allow for at least five general accounts of how to formulate the view:

[14] I take Reed (2013) to endorse a version of the **Conjunction Account**. See also Littlejohn (2011).

[15] It's not clear that any authors actually endorse the **Ambiguity Account**. Worsnip (2015) comes the closest. He seems to distinguish between a "robust" vs. a "less ambitious" formulation of fallibilism, where the former is characterized by an M-predicate, while the latter is characterized by an E-predicate. However, he seems to recognize only the former as capturing what he describes as the "core insight" behind fallibilism (Worsnip (2015, 225)). So he is perhaps best understood as endorsing what I call **Modal Purism** below.

[16] Leite (2010) and Stanley (2005) also endorse **Evidence Purism**.

[17] Worsnip (2015) seems to endorse **Modal Purism**; see n. 15. Lewis (1996) is typically read this way as well. But it's not clear that Lewis intends to characterize fallibilism by means of natural language epistemic modals. He describes fallibilism as the view that knowledge is compatible with "uneliminated possibilities of error" (Lewis (1996, 549)). But here the word "possibilities" is simply functioning as a noun that denotes a set of *metaphysically* possible worlds—namely, those in which the putatively known proposition is false. If the subject cannot eliminate these possibilities, the *theorist* might describe these possibilities as "epistemic"—and Lewis does. But as I cautioned above, we must be careful to distinguish between theoretical talk of "epistemic possibility" and natural language talk of what is possible, in the epistemic sense. Lewis never characterizes fallibilism by means of natural language sentences containing the words "possible" or "might".

The Equivalence Account: Fallibilism is a thesis about a knower's epistemic position and the truth-conditions of epistemic modals—but the two come to the same thing.

The Conjunction Account: Fallibilism is a thesis about a knower's epistemic position and the truth-conditions of epistemic modals—but the two do *not* come to the same thing.

The Ambiguity Account: "Fallibilism" is ambiguous between a thesis about (i) a knower's epistemic position, and (ii) a thesis about the truth-conditions of epistemic modals.

Evidence Purism: Fallibilism is just a thesis about a knower's epistemic position.

Modal Purism: Fallibilism is just a thesis about the truth-conditions of epistemic modals.

I argue below in favor of **Evidence Purism**.

4.3 THE SEMANTICS OF EPISTEMIC MODALS

I will argue for **Evidence Purism** by first setting out the problems facing the competing accounts. All of these problems have a common source: namely, an erroneous picture of the semantics of epistemic modals.

It is widely held that epistemic modals have a kind of mentalistic semantics on which they describe what people know, or what people know with certainty, or the state of their evidence, and so on.[18] In general terms: it is thought that *epistemic modals describe epistemic states*. My aim in this section is to show that this semantics is false: epistemic modals have no default function where they describe anyone's epistemic state. In the next section I draw out the consequences of this point, showing how it undermines all of the competing accounts of fallibilism except for **Evidence Purism**.

It is not surprising that the literature on fallibilism is dominated by the assumption that epistemic modals have a mentalistic semantics. This type of semantics is arguably the

[18]See the references in n. 4 for examples of such theories.

orthodox view of the matter, going back at least to G.E. Moore. Here is how Moore states the view:

> People *in philosophy* say: The propositions that I'm not sitting down now, that I'm not male, that I'm dead, that I died before the murder of Julius Caesar, that I shall die before 12 tonight are 'logically possible'. But it's not English to say, with this meaning: It's possible that I'm not sitting down now etc.—*this* only means 'It is not certain that I am' or 'I don't know that I am' (Moore (1962, 184); his emphasis).

Moore offers what has become known as a contextualist account of epistemic modals: *it's possible that* ϕ expresses different propositions on different occasions of use. If Bill uses this sentence, it expresses a proposition about what he knows; if Mary uses the sentence, it expresses a propostion about what she knows, and so on. Since Moore takes modalized sentences to express propositions about epistemic states, I will classify his theory as a version of what I will call **Mentalistic Contextualism**: the proposition expressed by a modalized sentence is true at a world w depending on how things stand with some contextually relevant subject or group's epistemic state at w. On such a theory, "possible" always means "possible for S", where S is the contextually relevant subject or group whose epistemic state is described by the proposition expressed by the sentence. As Stanley (2005, 127) puts it: "Since knowledge requires a knower, epistemic possibility does as well."

One might instead offer a mentalistic semantics in a relativist framework. On the latter, *it might be that* ϕ does not describe how things stand with any *particular* S's epistemic state. Rather, a use of the sentence expresses a general property an information state might have, and so the formal object expressed by the sentence is true relative to some information states and false relative to others. For example, on Egan's (2007) relativism, *it might be that* ϕ is true relative to a particular individual just in case the evidence within the individual's "epistemic reach" is compatible with ϕ. Egan's theory is a version of what I will call **Mentalistic Relativism**: the information state at which a modal claim is true is a state of information that is present in or somehow connected to the knowledge, certain knowledge, evidence, etc. of some individual or group.

We have, then, two different types of mentalistic semantics for epistemic modals—one based in a contextualist framework, the other based in a relativist framework. I will now argue that both are mistaken.

My argument draws on linguistic data first noted by Seth Yalcin (Yalcin (2007, 2011)). Yalcin's observations have been widely discussed in the larger literature on epistemic modals. But his work has yet to be discussed by epistemologists working on subjects informed by the semantics of epistemic modals. As we'll see, Yalcin's data is of great significance for how to theorize about fallibilism. But the true strength of his observations will only become clear when juxtaposed with some additional data I introduce below.

Yalcin's key observation is that there is a difference between Moore-paradoxical sentences like *ϕ but I don't know that ϕ* and what Yalcin calls *epistemic contradictions*—e.g. *ϕ but it might be that not-ϕ*. Consider:

(3) # It's raining, but I don't know it's raining.[19]

(4) # It's raining, but it might not be raining.

(5) Suppose it's raining, but I don't know it's raining.

(6) # Suppose it's raining, but it might not be raining.

As (3) and (4) illustrate, Moore-paradoxical sentences and epistemic contradictions are both bad things to assert. However, (5) and (6) show that there is a clear contrast in the embedding behavior of these sentences: it's perfectly coherent to merely *entertain* a Moore-paradoxical sentence, while epistemic contradictions continue to sound defective even in suppositional environments.

Nothing turns on the fact that (3) and (5) describe the speaker's knowledge. The contrast persists even if we vary the type of epistemic state or the subject whose epistemic state it is:

(6) # Suppose it's raining, but it might not be raining.

(7) Suppose it's raining, but I don't have any evidence for thinking that it's raining.

(8) Suppose it's raining, but I'm not certain that it's raining.

(9) Suppose it's raining, but none of the forecasters knows that it's raining.

[19] "#" marks sentences that are grammatical but infelicitous.

All of this constitutes a straightforward problem for **Mentalistic Contextualism**: if epistemic modals express propositions about epistemic states, then why is there a clear contrast in the embeddability of epistemic contradictions and Moore-paradoxical sentences?

The data is also a problem for **Mentalistic Relativism**, although this point has not been discussed in the literature. Recall Egan's semantics from above: *it might be that* ϕ is true relative to a particular individual just in case the evidence within the individual's "epistemic reach" is compatible with ϕ. Consider, then, the state of mind of believing *it might be that* ϕ: to believe this claim is to believe that the evidence within one's epistemic reach is compatible with ϕ.[20] Similarly, to suppose *it might be that* ϕ is to suppose that the evidence within one's epistemic reach is compatible with ϕ. But then why is it incoherent to suppose an epistemic contradiction? On Egan's semantics, this state of mind would just be one in which one supposes that ϕ is true but that the evidence within one's epistemic reach is compatible with not-ϕ.[21]

The recent literature contains several attempts to accommodate Yalcin's data in the framework of a mentalistic semantics.[22] Here are the three main strategies:

(i) Challenge the data.

(ii) Claim that the data results from the special contextual flexibility of epistemic modals.

(iii) Claim that the data results from general-purpose domain restriction mechanisms.

I take each strategy in turn.

On the first, some authors claim we can get a felicitous reading of embedded epistemic contradictions with the right contextual setup, or perhaps by reversing the order of the conjuncts. I have no qualm with these points. But they don't address the original data: Why do we need any special setup to get a felicitous reading in the first place? No such setup is needed to access a felicitous reading of embedded Moore-paradoxical sentences. Dowell (ms.) objects that the original data is unreliable because we haven't been given any context for interpreting the sentences. But in fact, the effect persists even if we do add some context:

[20] Egan himself is explicit on this point. See Egan (2007, 9).

[21] The data is also a problem for MacFarlane's relativism, but it takes a bit more work to set this out. I return to this point below when discussing Yalcin's semantics. See n. 29.

[22] See Crabill (2013), Dorr and Hawthorne (2013), Dowell (ms.), Ninan (forthcoming), Silk (2017), Sorensen (2009), Stojnić (2016, forthcoming), and Yanovich (2014).

John lives in a town where there has been an outbreak of a virus. Often times those infected with the virus initially show no symptoms.

(10) Imagine a scenario in which John is infected, but he doesn't know he's infected.

(11) # Imagine a scenario in which John is infected, but he might not be infected.

(12) # Imagine a scenario in which John is infected, but he is not infected.

There is a clear contrast between the acceptibility of (10) and (11). So denying the data is not a promising strategy.

On strategies two and three, one concedes that embedded epistemic contradictions sound bad and that this reveals something important about the semantics of epistemic modals. But one attempts to explain the data in the context of a mentalistic semantics.

On strategy two, one attempts to explain the data by the contextual flexibility of epistemic modals. For example, some claim that epistemic modals exhibit such a great deal of contextual flexibility that it's hard to understand whose epistemic state is at issue in suppositional contexts.[23] It is argued that epistemic modals can characterize the epistemic state of the speaker, some other subject, or even some relevant group. It is then claimed that assertions of modalized sentence perhaps default to describing the speaker's epistemic state—but not suppositions. After all, we have wide latitude in what we chose to merely *entertain*. So if we're not given any contextual pointers for whose epistemic state matters, this may explain why embedded epistemic contradictions behave differently than explicit descriptions of people's epistemic states, as in embedded Moore-paradoxical sentences.

Strategy three works in the opposition direction. Here one claims that the modal does receive a clear reading. But one argues that this reading makes the sentence as a whole express a literal contradiction. To see how this explanation works, note that modals are typically understood as quantifiers whose domains are restricted by contextually relevant propositions: *it might be that* ϕ is true at a context just in case the set of worlds compatible with the contextually relevant propositions contains at least one world at which ϕ is true.[24] Now recall the form of an epistemic contradiction: ϕ *but might not-*ϕ. If the proposition

[23] See Crabill (2013) and Dorr and Hawthorne (2013). Dorr and Hawthorne's (2013) account has various additional wrinkles that I won't be able to address here. See Chapter 3, Marushak and Shaw (ms.), and Silk (2017) for discussion of the problems facing their view.

[24] See Kratzer (1977, 1981, 2012) for the canonical semantics.

expressed by ϕ restricts the domain of the modal, then there will be no not-ϕ worlds in the modal's domain, and so the second conjunct—and hence the entire sentence—is guaranteed to be false. Authors defending this explanation typically rely on general-purpose domain restriction effects to account for why this restriction effect occurs in the case of epistemic contradictions.[25] For instance, Silk (2017) and Stojnić (2016, forthcoming) both claim that the effect essentially results from a phenomenon known as "local accommodation", wherein material later in a sentence is interpreted in light of material earlier in the sentence. So even if the modal by default describes some epistemic state, the first conjunct adds a further proposition that renders the sentence contradictory. No such effect is possible in the case of a Moore-paradoxical sentence. Here the second conjunct is just an explicit description of what someone knows. And the latter has no domain that can be restricted by material earlier in the sentence.

Strategies two and three succeed in identifying mechanisms that could account for Yalcin's data. But this virtue turns out to be a vice. Consider:

(13) All of the evidence points to Smith.

This sentence is context-sensitive: it expresses different propositions on different occasions of use. And the contextually flexibility is wide indeed. If uttered by a juror, (13) might express the proposition that the evidence presented to the jury points to Smith. If uttered by a detective, (13) might express the proposition that some larger body of evidence, available only to the detectives, points to Smith. And if Smith is on trial for two different crimes, uses of (13) might express propositions about completely unrelated bodies of evidence.

Note also: to assert (13) is to describe how things stand with someone's epistemic state. Evidence itself may consist of propositions, or even physical stuff that can be placed in an evidence locker. But when one uses (13) in assertion, one describes the evidence someone *has* or *possesses*. It is universally agreed that this sort of "having" or "possession" relation is somehow epistemic: it concerns some sort of access the subject enjoys to the evidence itself.[26] Thus, to use (13) is to describe how things stand with someone's epistemic state.

[25] Dorr and Hawthorne (2013) is the exception. Again, see Chapter 3, Marushak and Shaw (ms.), and Silk (2017) for discussion.

[26] See Comesaña and McGrath (2014), Henning (2014), Hornsby (2008), Lord (2010, 2015), Neta (2008*b*, 2011), Schroeder (2008, 2011), and Skorupski (2010, 108).

Now, consider how (13) embeds:

(14) Suppose Jones is the murderer, but all of the evidence points to Smith.

This sentence sounds perfectly fine. It is easy to understand what it means: we are asked to imagine a scenario in which Jones is actually the murderer but the evidence available to some individual or group is misleading and points to Smith. That is, (13) continues to describe epistemic states *even when embedded*. By hypothesis, Jones is the murderer, but *this* piece of information is not counted among the relevant evidence even though we are quantifying over "all" of the evidence and the quantifer is placed immediately after the sentence describing Jones as the murderer. What this shows is that there is a clear default for reading talk of "all of the evidence" as describing the evidence people "have" or "possess". To merely suppose that Jones is the murderer is not to suppose that anyone knows this or has any evidence for thinking it's true. This is why Jones's being the murderer does not figure as part of the relevant evidence described in the second conjunct.

Return now to our second and third strategies for explaining the epistemic contradiction data. Recall that these strategies appeal to the following claims: embedded epistemic contradictions sound bad because (a) it's not clear whose epistemic state is at issue; (b) the first conjunct restricts the domain of the modal in the second.

However, (14) runs counter to the predictions of each of these theories. Regarding (a), talk of "all of the evidence" exhibits wide contextual flexibility, this language continues to describe epistemic states even when embedded, and we are given no contextual cues for whose epistemic state is at issue. But (14) sounds perfectly fine.

Regarding (b), talk of "all of the evidence" has a domain that can be restricted by information described earlier in a sentence. "all of the evidence" literally contains a quantifier that is ranging over bodies of information—here, bodies of evidence. And yet, the first conjunct of (14) does not restrict the domain of the quantifier in the second. That is, (14) does not ask us to entertain a scenario in which Jones is the murderer but all of the evidence—*including the fact that Jones is the murderer*—points to Smith.

Thus, we have a dilemma: strategies two and three either generate false predictions, or they fail to explain Yalcin's data in the first place. That is, the theories as stated predict

that (14) should sound bad, but it doesn't. However, if these theories are adjusted to block this prediction, then it's not clear how they can explain the infelicity of embedded epistemic contradictions in the first place. After all, evidence claims have exactly the sort of semantic profile that a mentalistic semantics attributes to modal claims. So it would seem that one either predicts that both are fine to embed, or both are not. But neither outcome is correct.

Importantly, this isn't some isolated problem facing only strategies two and three. This is a general problem facing *any* attempt to give a pragmatic explanation of the epistemic contradiction data within the framework of a mentalistic semantics. If epistemic modals have the same semantic profile as talk of "all of the evidence", then it's hard to see how any pragmatics could explain the difference in their embedding behavior. Both expressions will be contextually flexible characterizations of epistemic states. So it's hard to see what sort of pragmatics could possibly engage with one type of language but not the other so as to predict their difference in embedding behavior.[27]

All of this seems to show that epistemic modals are not in the business of describing epistemic states. There are many expressions in our language that clearly have this function. But these expressions do not embed like epistemic modals.

What, then, is the semantics of epistemic modals? I will briefly sketch two proposals that avoid the problems just raised for mentalistic semantics. The first is the proposal that Yalcin himself advances to explain the data. Yalcin's is an *expressivist* semantics: epistemic modals have no descriptive function at all.[28] They simply express features of a speaker's state of mind. For example, to assert "it might be that ϕ" is to express the compatibility of ϕ with one's belief state as a whole. Similarly, to believe a modal claim is not to believe some proposition about anyone's epistemic state—it's just a matter of *being* in a doxastic state whose content is compatible with ϕ.

This account of the state of mind of accepting an epistemic modal explains the infelicity of supposing ϕ *but might not-*ϕ. To suppose that ϕ is to get into an information state that contains only ϕ-worlds. But to suppose that *might not-*ϕ is to get into an information state that contains at least some not-ϕ worlds. No coherent information state has both properties,

[27] See Chapter 3 for discussion of an analogous problem concerning the language of epistemic reasons.

[28] For other examples of expressivism about epistemic modals, see Moss (2013, 2015) Rothschild (2012), and Swanson (2006, 2011).

which explains why it sounds odd to request that one enter such a state.[29] Importantly, it is open for Yalcin and other expressivists that hold that talk of "all of the evidence" does *not* have a similar expressivist semantics. This semantic difference can then explain the contrast in embedding data.[30]

An alternative proposal is what I will call *veritic semantics*: epistemic modals have a descriptive function, but the information they describe is not restricted to information connected to anyone's epistemic state. On this semantics, epistemic modals simply characterize relevant *truths*.[31] For example, "it might rain tonight" may simply describe the relevant meteorological facts as being compatible with rain. The sentence therefore expresses a proposition that is true at a world depending on how things stand with the meteorological facts at that world. Of course, it may turn out that the speaker knows the relevant facts at the world of the context of use—indeed, it may be inappropriate for the speaker to assert the sentence *unless* she knows the relevant facts. But this does not mean that the sentence describes the speaker's knowledge: there will be worlds at which the speaker fails to know the relevant facts, but the truth of the modal at such worlds still just turns on whether rain is compatible with the meteorological facts at those worlds. Compare a sentence like *It's raining*: it may be inappropriate for a speaker to assert this sentence unless she knows it's raining, but this does not mean that the sentence itself describes how things stand with the speaker's knowledge.

Veritic semantics offers a different explanation of the epistemic contradiction data. Recall the form of an epistemic contradiction: ϕ *but might not-*ϕ. To suppose that ϕ is to suppose that ϕ is true. But the modal in the second conjunct of an epistemic contradiction

[29]Herein lies the problem for MacFarlane's version of **Mentalistic Relativism**. MacFarlane and Yalcin endorse the same compositional semantics for epistemic modals: both take *might* ϕ to express a property of an information state s that is true of s iff s is compatible with ϕ. But MacFarlane explicitly rejects Yalcin's account of the state of mind of accepting *might* ϕ (see MacFarlane (2014, 278–279)). MacFarlane takes it to be a problem for Yalcin that the state of believing an epistemic contradiction is "conceptually impossible", as MacFarlane puts it (again, see MacFarlane (2014, 278–279)). He wishes to allow for the possibility of such that states of belief while deeming them irrational, and he appeals to norms on belief to secure the latter judgment. But what norms on supposition make it irrational to merely *suppose* an epistemic contradiction? On MacFarlane's own semantics, to suppose such the latter should just be to entertain a context of assessment wherein ϕ is true but the contextually relevant state of knowledge is compatible with not-ϕ. So MacFarlane can't explain the epistemic contradiction data in the first place.

[30]However, see Chapter 3 for some larger problems with this strategy.

[31]See Marushak and Shaw (ms.) and Chapter 3 for development and defense of this theory.

is characterizing how things stand with relevant truths. We should therefore expect the proposition that ϕ to count as part of the information characterized by the modal. But then the second conjunct comes out false, for the relevant information cannot be compatible with not-ϕ if the relevant information contains the information that ϕ.

Veritic semantics exploits a simple parallel with other cases of domain restriction. Consider:

(15) Suppose John's marble is blue, but every marble is red.

This request strikes us as incoherent, since there can be no such scenario. The first conjunct restricts the domain of the quantifier in the second: "every marble" includes John's marble, so any scenario in which the first conjunct is true is a scenario in which the second conjunct is false. Notice that these domain restriction effects hold only at a context in which the sentences are both used together. There are many contexts where "every marble is red" is true, but "John's marble is blue" is also true because the John's marble is not salient in the context and so fails to be included in the quantifier domain. In other words, there is no *semantic* defect in sentences like "John's marble is blue, but every marble is red". The defect is pragmatic and concerns the fact that the entire sentence typically expresses a contradiction at a given context of use.

The same goes for epistemic contradictions, on veritic semantics. The modal is quantifying over relevant truths, and the first conjunct makes ϕ is salient as being true, so ϕ restricts the domain of the subsequent modal. Again, the domain restriction effect occurs only at a context in which the sentence is used. There are many contexts at which "it might be that not-ϕ" and ϕ are both true since ϕ need not restrict the domain of the modal unless it is *salient* as being true.

This semantics can also account for the difference in the embedding behavior of evidence-talk. Recall our example from above:

(16) Suppose Jones is the murderer, but all of the evidence points to Smith.

If talk of "all of the evidence" is by default functioning to describe bodies of possessed evidence, then the information that Jones is the murderer will not on its own restrict the domain of the quantifier. So the difference between the embedding behavior of modal talk

and evidence talk lies in the fact that the former by default characterizes relevant *truths*, while the latter by default characterizes relevant *epistemic states*.[32] The key point here is that the veritic explanation of the epistemic contradiction data does not rely on *general purpose* domain restriction mechanisms. The explanation relies on the fact that modals are by default characterizing mere truths. So if evidence talk does not have such a semantics, it's not surprising that it fails to display the embedding behavior of modal talk.

4.4 FALLIBILISM AND EPISTEMIC MODALS

I've just argued that epistemic modals have no default function where they describe how things stand with anyone's epistemic state. And I set out two alternative semantics: a veritic semantics, on which epistemic modals describe how things stand with facts that need not be related to anyone's epistemic state, and an expressivist semantics, on which epistemic modals perform no descriptive function at all. I will now apply these semantics to the question of how one ought to formulate fallibilism.

4.4.1 Problems for the Equivalence Account

It is an immediate consequence of the semantics I've offerred that the **Equivalence Account** is false. On this account, fallibilism can be equivalently characterized as both a thesis about the type of epistemic position consistent with knowledge and the semantics of epistemic modals, on their default use. But the default use of epistemic modals does not function to express propositions about how things stand with anyone's epistemic position. So claims about the semantics of epistemic modals are not equivalent to claims about the type of epistemic position consistent with knowledge.

For example, suppose one wishes to characterize a fallibilist theory as one on which CKAs are sometimes true. Veritic semantics will then count as a fallibilist theory, for this semantics entails no *semantic* defect in CKAs. Instead, the defect in typical CKAs is pragmatic: their

[32]See Chapter 3 for further discussion.

infelicity results from the fact that CKAs entail epistemic contradictions, and familiar domain restriction effects tend to make the latter express contradictions.[33] A true CKA requires only that there be a context in which the putatively known proposition fails to restrict the domain of the modal. It may prove difficult to construct such a context, but nothing in our semantics rules this out. Suppose, then, that there is such a context in which a CKA is true. What follows about the type of epistemic position consistent with knowledge? Answer: nothing. A context in which a CKA is true will be one in which S knows that ϕ but it is compatible with the relevant truths that not-ϕ. But nothing in the semantics forces these relevant truths to concern how things stand with S's epistemic position. So nothing follows about whether the context is one in which S has fallible knowledge, in the **Evidence Gloss** sense of the phrase.

4.4.2 Problems for Modal Purism and the Ambiguity Account

According to **Modal Purism** and the **Ambiguity Account**, fallibilism—or at least one sense of "fallibilism"—ought to be characterized in exclusively modal terms. This is an odd proposal. Consider Worsnip's characterization of fallibilism quoted above: "S can sometimes truly assert 'it is possible that not-p' even though S knows that p" (Worsnip (2015, 226)). On this formulation, fallibilism is simply equivalent to the negation of what Egan, Hawthorne and Weatherson (2005) call the "speaker inclusion constraint": "whenever S truly utters a *might be* F, S does not know that a is not F" (Egan, Hawthorne and Weatherson (2005, 135)). Egan, Hawthorne, and Weatherson themselves argue against this constraint—and hence in favor of fallibilism, on Worsnip's characterization of the view. Here is the type of example they press against the constraint:

> Tom is stuck in a maze. Sally knows the way out, and knows she knows this, but doesn't want to tell Tom. Tom asks whether the exit is to the left. Sally says, 'It might be. It might not be.' Sally might be being unhelpful here, but it isn't clear that she is lying. Yet if the speaker-inclusion constraint applies to unembedded epistemic modals, then Sally is

[33] Worsnip (2015) offers a similar account of the infelicity of CKAs. I explain the key difference below in §4.5.

> clearly saying something that she knows to be false, for she knows that she knows which way is out (Egan, Hawthorne and Weatherson (2005, 140)).

Now, Egan, Hawthorne, and Weatherson are themselves cautious about just how much this example shows. But suppose the case works and falsifies the speaker inclusion constraint. Would anyone take this to vindicate fallibilism in any recognizable sense of the term? Obviously not. All we would have learned is that speakers can use epistemic modals to characterize the limited information available to some poorly informed third party. But fallibilism is manifestly not a view about *that*.

Perhaps we need to try characterizing fallibilism by a different version of the **Modal Gloss**: suppose fallibilism *just is* the view that CKAs are sometimes true, where this is not equivalent to any claim about how things stand with anyone's epistemic position. As noted above, our veritic semantics will count as a fallibilist theory on this construal. But notice: expressivism will not. For one, the notion of truth is not entirely applicable to uses of epistemic modals on the expressivist picture, as modals do not serve to describe what the world is like. More importantly, expressivism is a view on which CKAs *do* suffer from a kind of semantic defect—namely, the same defect that afflicts epistemic contradictions. We reviewed this defect above when stating Yalcin's explanation of the epistemic contradiction data: no coherent information state can accept both conjuncts of an epistemic contradiction. And since CKAs entail epistemic contradictions, the same goes for CKAs. Here, then, is the upshot of our proposed characterization of fallibilism: veritic semantics is a fallibilist theory, but expressivism is not.

Is this a reasonable characterization of the fallibilist thesis, on any sense of the term? Again, obviously not. The difference between these semantics is an important issue in the philosophy of language. But these issues would appear to have nothing to do with the nature of knowledge, skepticism, and so on. To be clear: I am not claiming that issue of which semantics is correct will have *no* interesting downstream epistemological consequences. Perhaps it does. The point is simply that it's difficult to see how any of this has anything to do with what made epistemologists interested in *fallibilism* in the first place.

For example, fallibilism is typically thought to block skeptical arguments like the following:

P1. Knowledge requires entailing evidence.

P2. No one has entailing evidence for anything.

C. No one knows anything.

In particular, fallibilism is typically thought to block P1. But fallibilism *can't* block P1 unless it is at least in part a thesis about the type of epistemic position consistent with knowledge. Here, then, is another way of putting the argument in this section: fallibilism is widely thought to be one of the main strategies for responding to skepticism. But it's not obvious how a very subtle point in semantics—i.e. whether CKAs are semantically defective, or only contradictory at typical contexts—could capture what epistemologists have been trying to get at by developing a fallibilist response to skepticism.

4.4.3 Problems for the Conjunction Account

Considerations like the above might lead one to favor the **Conjunction Account**. Recall that on this account, fallibilism is a thesis about a knower's epistemic position and the truth-conditions of epistemic modals—but the two do not come to the same thing. This account might appear to offer the best of both worlds: we maintain a connection between fallibilism and ordinary talk of possibility, while at the same time building in some properly epistemological content that serves our anti-skeptical ends.

To be sure, the **Conjunction Account** makes fallibilism an epistemologically important view. But there is now a different problem: by building in a modal element, the view makes the truth of fallibilism *turn* on irrelevant matters. Recall the point from above: veritic semantics will count as a fallibilist theory on a common modal characterization, while expressivism will not. Suppose expressivism is correct. Fallibilism itself then turns out to be false on the **Conjunction Account**, since the view entails a claim about the semantic consistency of CKAs (say). This is an odd result. As we discussed above, the difference between each semantics has nothing to do with the nature of knowledge or skepticism. So the issue of which is true should not decide the fate of fallibilism itself. In other words, the **Conjunction Account** may appear to be a weakening of the problematic claims we saw with the **Modal Account**. But the problems just re-appear, since the **Conjunction**

Account still leaves fallibilism with *commitments* about the semantics of epistemic modals, even if these commitments do not exhaust the content of the view.

4.4.4 In Favor of Evidence Purism

We've arrived at the view that I favor. The advantages of **Evidence Purism** are straightforward. On this account, fallibilism is clearly an epistemologically significant thesis. The view is a claim about the type of epistemic position consistent with knowledge—e.g. that knowledge can be based on non-entailing evidence. Such a view has obvious anti-skeptical purport, so it matters greatly whether this view is true. In addition, **Evidence Purism** presupposes the correct semantics for epistemic modals: there is no default use of these modals that characterizes epistemic states, so there can be no equivalent **Modal Gloss** formulation of fallibilism. And by separating the fallibilist thesis from any commitments about epistemic modals, we avoid all of the problems facing the other formulations.

4.5 HOW NOT TO ARGUE ABOUT FALLIBILISM

Getting clearer on the semantics of epistemic modals helps us understand what fallibilism *is*. But we also learn things about how to argue about whether the view is *true*. In particular, we learn that many prominent arguments—both for and against fallibilism—fail to hit their mark.

I begin with what is perhaps the most widely discussed argument against the fallibilism: namely, the argument from the infelicity of CKAs.[34] As we noted above, CKAs typically sound odd and contradictory: e.g. "I know it's raining, but it might not be raining." This is supposed to be a problem for fallibilism because CKAs plausibly describe instances of fallible knowledge.

[34]This argument is typically attributed to Lewis (1996), but I am skeptical for the reasons set out in n. 17. For discussion of the CKA objection to fallibilism, see Anderson (2014), Dodd (2010), Dougherty and Rysiew (2009), Fantl and McGrath (2009), Leite (2010), Littlejohn (2011), Reed (2013), Rysiew (2001), Stanley (2005), and Worsnip (2015).

If fallibilism is at least in part the thesis *that CKAs are sometimes true*, then this is at least a good *prima facie* objection. But I've already argued against this conception of fallibilism above. Suppose instead that fallibilism is a thesis about the epistemic position consistent with knowledge. Now the CKA objection fails to hit the mark: the characteristic function of epistemic modals is not to describe anyone's epistemic position or other epistemic state. On the semantics I've defended in this paper, typical CKAs just express ordinary contradictions because the first conjunct restricts the domain of the modal in the second (both veritic semantics and expressivism yield this result, although by different means). So typical CKAs do not describe instances of fallible knowledge.

Worsnip (2015) offers a similar reply on behalf of the fallibilist. He claims that assertions update subsequent modal bases, and so again, the first conjunct of a CKA restricts the domain of the modal in the second conjunct and the sentence expresses a contradiction. However, there is a key advantage to the veritic and expressivist explanations: neither relies on features of the speech act of assertion.[35] Consider:

(17) # Suppose I know it's raining, but it might not be raining.

This sentence sounds just as bad as an unembedded CKA, even though neither conjunct is asserted.[36] A simple explanation is the following: the defect of typical CKAs ultimately stems from the fact that they entail epistemic contradictions. Once we have an explanation of the infelicity of embedded epistemic contradictions, we can capture the CKA data in its full generality.

Let me turn now to another recent argument against fallibilism. Dodd (2011) offers what appears to be a straightforward derivation of the claim that knowledge requires epistemic probability 1, where epistemic probability is the probability of a proposition given one's epistemic position. His argument has two premises:

P1. ϕ is epistemically possible for S only if S doesn't know the contradictory of ϕ to be true.

[35]This is also an advantage in comparison to the pragmatic explanations offerred by Dougherty and Rysiew (2009), Fantl and McGrath (2009), Littlejohn (2011), and Rysiew (2001). All appeal to Gricean pragmatics, which has no easy application in suppositional environments.

[36]Note that Worsnip cannot simply reply that suppositions update modal bases. Were he to make this move, he would face the same problem we raised in §4.3 against appealing to general-purpose domain restriction mechanisms to explain the infelicity of embedded epistemic contradictions.

P2. If the epistemic probability of ϕ on S's epistemic position is greater than 0, then ϕ is epistemically possible for S.

C. If S knows that ϕ, then the epistemic probability of ϕ on S's epistemic position equals 1.

This is a valid argument. By P2 and P1, we have the following:

P3. If the epistemic probability of ϕ on S's epistemic position is greater than 0, then S doesn't know the contradictory of ϕ to be true.

By contraposition, we have the following:

P4. If S knows that not-ϕ, then the epistemic probability of ϕ on S's epistemic position is not greater than 0.

The probability calculus then gives us this:

P5. If S knows that not-ϕ, then the epistemic probability of ϕ on S's epistemic position is equal to 0.

From which we conclude the following, again by the probability calculus:

C. If S knows that not-ϕ, then the epistemic probability of not-ϕ on S's epistemic position equals 1.

Let us first consider P1, repeated here:

P1. ϕ is epistemically possible for S only if S doesn't know the contradictory of ϕ to be true.

This premise contains the troublesome phrase "epistemically possible for S". Dodd is explicit that this is a theoretical term, but he argues for P1 by means of intuitions about natural language talk of possibility. Thus, if "epistemic possibility for S" is answerable to our ordinary use of epistemic modals, P1 is false: as I've already argued, epistemic modals do not have truth-conditions that concern how things stand with what people know.

Suppose we instead read P1 as a claim about some thing—epistemic possibility—that needn't line up with how we talk about possibility in natural language. There is clearly a

sense of epistemic possibility that concerns what is compatible with a subject's knowledge. So P1 is plausible on this reading. However, P2 now becomes question-begging:

P2. If the epistemic probability of ϕ on S's epistemic position is greater than 0, then ϕ is epistemically possible for S.

On the proposed reading of "epistemic possibility", P2 claims that if ϕ has non-zero epistemic probability, then ϕ is compatible with what S knows. But why should a fallibilist grant this premise? Many fallibilists literally employ a version of the **Evidence Gloss** on which fallible knowledge is consistent with a merely high epistemic probability of truth.[37] So these authors will all claim that one's knowledge rules out many propositions that have non-zero epistemic probability. Thus, Dodd faces a dilemma: his argument is either unsound or question-begging.

I turn finally to Worsnip's (2015, 238–239) recent argument in favor of fallibilism. He claims that the felicity of dialogues like the following suggests that fallibilism is true:

A: Do you know what the capital of South Africa is?

B: Yes, I think I know the answer to your question—Pretoria. But it might be Johannesburg.[38]

Worsnip's thought is this: suppose it is true in the imagined scenario that B knows the capital is Pretoria. We would then have a case where S knows that ϕ but S can truly assert that it is possible that not-ϕ. Hence, there exists fallible knowledge.

However, we've already seen in §4.4.2 that this is a bizarre characterization of fallible knowledge. Recall the example discussed in that section: suppose I know how to get out of a maze but I don't want to tell you, and so I say "The exit might be to the left, and it might to the right". Even if we could establish that this was a genuine case of truly asserting the possibility of not-ϕ while knowing that ϕ, no one would conclude that we had established anything of epistemological significance.

Worsnip's only recourse is to argue that the above dialogue contains some special use of an epistemic modal on which it describes B's epistemic position. We would then have

[37] See, for instance, Dougherty (2011), Dougherty and Rysiew (2009), Fantl and McGrath (2009), and Reed (2012).

[38] This example comes from Worsnip (2015, 232).

a case where B knows that ϕ but B's evidence (say) leaves open the possibility of not-ϕ. However, it's not clear why we should grant that the modal receives such a reading in this case. In saying "it might be Johannesburg", B might just be asserting the following proposition: "given my track record in geography, the answer might be Johannesburg." We can even concede that B describes some *aspect* of her epistemic position: "given my feeling of uncertainty, the answer might be Johannesburg." But all will concede that one can know that ϕ even if some limited aspect of one's epistemic position fails to guarantee that ϕ is true. What matters is whether knowledge is consistent with one's *entire* epistemic position failing to guarantee that ϕ is true. Thus, without further argumentation, Worsnip's example shows nothing of epistemological significance.

In sum: fallibilism is neither in whole nor in part a thesis about the truth-conditions of epistemic modals. Fallibilism is an epistemologically interesting thesis about the type of epistemic position consistent with knowledge. Epistemic modals have no default function where they describe anyone's epistemic position. So there is no straightforward way to rely on data about epistemic modals to help us determine whether fallibilism is true.

4.6 CONCLUSION

I've argued that the methodology that presently dominates the literature on fallibilism cannot succeed in helping us determine whether the view is actually true. How, then, should we proceed?

The most conservative strategy is to continue relying on data about epistemic modals, while exercising greater care to force readings where they target someone's epistemic position. This is not an easy task. As we saw above in connection with Worsnip's case, there are many readings of modals—even mentalistic readings—that fail to precisely target a subject's epistemic position. Of course, one could always force the desired reading by a restrictor phrase:

(18) I know the animal is a zebra, but given only the evidence on the basis of which I know this fact, it might be a painted mule.

But no one has intuitions about sentences like these—or rather, the intuitions one does have are liable to simply reflect one's antecedent theoretical committments for or against fallibilism.

An alternative strategy would be to focus on examples that do not contain epistemic modals at all. Fallibilism is sometimes thought to face problems from Gettier cases and lottery cases.[39] With regard to the former, it might appear that fallibilists are inevitably stuck with the view that knowledge requires a kind of luck, in which the putatively known proposition turns out to be true even though one's epistemic position does not guarantee its truth. With regard to the latter, many fallibilists are explicitly committed to the possibility of knowledge on the basis of merely high probabilities.[40] But this view generates a variety of puzzles concerning one's inability to know either the winner, or the losers, of an arbitrarily large lottery. These arguments have received less attention than the ones we've discussed concerning epistemic modals. So a re-evaluation is in order.

But however we choose to proceed, we must take care to ensure that the fallibilist thesis under discussion is indeed of epistemological significance. Only then can we be sure that our investigation is not in vain.

[39]See Reed (2012) for an overview of this literature.
[40]See the references in n. 37.

5.0 PROBABILITY MODALS AND INFINITE DOMAINS

5.1 INTRODUCTION

Natural language talk of what is *likely* or what is *probable* has a foot in two worlds: this language would appear to have ties to the mathematical theory of probability, and yet talk of what is likely is commonplace even among speakers with no special mathematical training. How then should we approach the semantics of this language?

Semanticists were initially reluctant to base their theories on the quantitative notion of probability employed by scientists and mathematicians. This latter notion may well have its origins in the folk concept of probability, but it was thought to be a mistake to credit ordinary language users with a tacit grasp of the mature theory itself.[1]

However, there has recently been a striking shift in opinion: semantics based on Kolmogorovian probability (hereafter: *probabilistic semantics*) have become widespread.[2] One reason for this change in attitude has been the recognition that non-probabilistic semantics cannot easily account for explicitly quantitative assessments of probability, as in *there is a 70% chance of rain*. But what is perhaps more surprising is that non-probabilistic semantics fail to capture intuitions about even basic inferences involving judgments of comparative likelihood.

For example, on Kratzer's canonical *world-ordering semantics* (Kratzer (1991*b*, 2012)), an ordering on worlds generates an ordering on propositions, which in turn fixes the facts

[1] See Hamblin (1959, 234), Koopman (1940, 269-270), and Kratzer (2012, 25) for expressions of this sentiment.

[2] Defenders include Cariani (2016), Carr (2015), Lassiter (2010, 2011, 2015), Moss (2013, 2015), Rothschild (2012), Swanson (2006, 2011, 2016), and Yalcin (2007, 2010, 2011).

about what is more/less/equally likely than what.[3] But Lassiter (2010, 2011, 2015) and Yalcin (2010) demonstrate that this semantics validates clearly invalid inference patterns like the following:

> The coin's landing heads is as likely as its landing tails. Therefore, the coin's landing heads is as likely as any proposition whatsoever.

In contrast, probabilistic semantics validates a variety of intuitively valid inference patterns and fails to validate obviously invalid inference patterns like the one above.[4]

Nevertheless, recent work by Holliday and Icard (2013) suggests that the shift away from world-ordering semantics may be premature: they present a world-ordering semantics that promises to capture the core inferences just as well. And indeed, their work has been widely thought to achieve this aim.[5]

In this paper, I argue that the challenge remains: defenders of world-ordering semantics have yet to offer a plausible semantics that captures the logic of comparative likelihood. I show that Holliday & Icard's semantics fails to validate one intuitively valid inference pattern that is validated by probabilistic semantics. The countermodel turns on distinctive features of infinitely large domains, but probabilistic semantics validates the inference nonetheless. I go on to consider several ways of patching Holliday & Icard's semantics to validate the desired inference. But I argue that each has considerable costs that have no analogue for probabilistic semantics. As a result, probabilistic semantics remains the better explanation of the data.

5.2 PROBABILISTIC VERSUS WORLD-ORDERING SEMANTICS

Let's begin by contrasting two approaches to the semantics of natural language probability talk. (Readers eager to see the countermodel can skip to §5.4.)

[3] The label *world-ordering semantics* is due to Holliday and Icard (2013).
[4] See Lassiter (2010, 2011, 2015) and Yalcin (2010).
[5] See Cariani (2016, n. 9), Carr (2015, 697 n. 23), Lassiter (2015, 663), and Suzuki (2013, 216).

Start with a standard propositional language $\mathscr{L}$ and extend it as follows:

- If ϕ and ψ are sentences of $\mathscr{L}$, then $\ulcorner(\phi \geqslant \psi)\urcorner$ is a sentence of $\mathscr{L}$.
- If ϕ is a sentence of $\mathscr{L}$, then $\ulcorner\Diamond\phi\urcorner$ is a sentence of $\mathscr{L}$.[6]

Sentences of the form $(\phi \geqslant \psi)$ are intended to model natural language judgments of comparative likelihood: *ϕ is at least as likely as ψ*. Sentences of the form $\Diamond\phi$ model talk of epistemic possibility: *it might be that ϕ*.

We also add the following definitions:

- $(\phi > \psi)$ models *ϕ is more likely than ψ* and is defined as $(\phi \geqslant \psi) \wedge \neg(\psi \geqslant \phi)$.
- $\Delta\phi$ models *it is likely that ϕ* and is defined as $(\phi > \neg\phi)$.
- $\Box\phi$ models talk of epistemic necessity—*it must be that ϕ*—and is defined as $\neg\Diamond\neg\phi$.

What, then, is the semantics appropriate to $\mathscr{L}$, in its intended interpretation? There are two main approaches, which diverge in their semantics for $\geqslant$. Before setting out these views, let's give the semantic entries they have in common.

Let a model be a tuple $\mathscr{M} = \langle W, S, V\rangle$, where W is a non-empty set (intuitively, a set of metaphysical possibilities), S is a non-empty set (intuitively, an information state or a set of epistemic possibilities), and V is a function assigning elements of $\mathscr{P}(W)$ to atomic sentences A=$\{p,q,...\}$ (intuitively, V specifies which proposition is expressed by a given atomic sentence). We then define an interpretation $[\![\cdot]\!]_{\mathscr{M}}$ for $\mathscr{M}$ as follows:

- $[\![\phi]\!]_{\mathscr{M}} = V(\phi)$ if $\phi \in A$.
- $[\![\neg\phi]\!]_{\mathscr{M}} = W - [\![\phi]\!]_{\mathscr{M}}$.
- $[\![(\phi \wedge \psi)]\!]_{\mathscr{M}} = [\![\phi]\!]_{\mathscr{M}} \cap [\![\psi]\!]_{\mathscr{M}}$.
- $[\![\Diamond\phi]\!]_{\mathscr{M}} = \{w \in W : [\![\phi]\!]_{\mathscr{M}} \cap S \neq \emptyset\}$.[7]

A sentence ϕ is true at w in $\mathscr{M}$ ($[\![\phi]\!]^w_{\mathscr{M}} = 1$) iff $w \in [\![\phi]\!]_{\mathscr{M}}$. A sentence is valid in $\mathscr{M}$ iff it is true at every $w \in W$ in $\mathscr{M}$. A sentence is valid in a class of models $\mathscr{C}$ iff it is valid in every

[6] I omit the relevant corner quotes from here on out.

[7] I follow MacFarlane (2011, 2014) and Yalcin (2007) in using S to determine the interpretation of $\Diamond\phi$, as opposed to an accessibility relation. I opt for the former approach to simplify the semantics, since the question of the world-sensitivity of $\Diamond\phi$ is not at issue in what follows.

model in the class. A semantics validates an inference pattern iff the material conditional whose antecedent is the conjunction of the premises and whose consequent is the conclusion is valid in the class of models defined by the semantics.

Our first approach to the semantics of $\geqslant$ treats judgments of comparative likelihood as qualitative comparisons of propositions, where these comparisons are in turn grounded in a more fundamental ranking of the worlds that comprise each proposition.

For example, on Kratzer's (1991*b*) semantics for $\geqslant$, context delivers a set of propositions O—called the *ordering source*—that induces a pre-order, $\succeq_O$, on the members of S as follows:[8]

$$w \succeq_O w' \text{ iff } \{\alpha \in O : w \in \alpha\} \supseteq \{\alpha \in O : w' \in \alpha\}.$$

Intuitively, the ordering source relevant to $\geqslant$ represents a standard of normality, and worlds are ranked higher the closer they come to matching the normal course of events.

Kratzer then uses this ranking of worlds to determine a ranking, $\gtrsim$, of propositions:

$$\alpha \gtrsim \beta \text{ iff } \forall w \in \beta : \exists w' \in \alpha : w' \succeq_O w.$$

That is, α is ranked at least as high as β iff every β-world can be paired with an α-world that is at least as highly ranked.[9] Finally, Kratzer takes $(\phi \geqslant \psi)$ to be true at a world in a model iff $[\![\phi]\!]_{\mathscr{M},S} \gtrsim [\![\psi]\!]_{\mathscr{M},S}$, where $[\![\phi]\!]_{\mathscr{M},S} = [\![\phi]\!]_{\mathscr{M}} \cap S$.

Generalizing from the particulars of Kratzer's approach, *world-ordering semantics* takes models to be the following: $\langle W, S, V, \succeq, \uparrow \rangle$, where $\succeq$ is a pre-order on S, and $\uparrow$ is a lifting operation—that is, a function from $\succeq$ to a binary relation, $\succeq^{\uparrow}$, on $\mathscr{P}(S)$. The role of the lifting operation is to take us from a ranking on worlds to a ranking on propositions. So, for example, Kratzer's definition of $\gtrsim$ is one way to lift a pre-order on worlds to a pre-order on propositions. We finally let $[\![(\phi \geqslant \psi)]\!]^w_{\mathscr{M}} = 1$ iff $[\![\phi]\!]_{\mathscr{M},S} \succeq^{\uparrow} [\![\psi]\!]_{\mathscr{M},S}$.[10]

[8] Note that a pre-order is a reflexive and transitive binary relation.

[9] This method of generating a ranking of propositions from a ranking of worlds is due to Lewis (1973).

[10] This generalization of Kratzer's semantics and the term *lifting operation* are due to Holliday and Icard (2013).

However, as Lassiter (2010, 2011, 2015) and Yalcin (2010) point out, world-ordering semantics with Kratzer's (1991*b*) lifting operation faces what Lassiter (2015) calls the *disjunction puzzle*—namely, the semantics validates the following inference pattern:

I1:

P1. $\phi \geqslant \psi$

P2. $\phi \geqslant \chi$

C. $\phi \geqslant (\psi \vee \chi)$

I1 is clearly invalid: from the fact that heads is at least as likely as heads, and heads is at least as likely as tails, it does not follow that heads is at least as likely as heads or tails.[11]

Lassiter and Yalcin use the disjunction puzzle to motivate an alternative semantics for $\geqslant$. On *probabilistic semantics*, judgments of comparative likelihood are grounded in a quantitative ranking of propositions. Models are as follows: $\langle W, S, V, \mathscr{F}, \mu \rangle$, where $\mathscr{F}$ is a σ-algebra of subsets of S, and μ is a finitely additive probability measure. That is, $\mathscr{F}$ is a subset of $\mathscr{P}(S)$ such that $S \in \mathscr{F}$, and $\mathscr{F}$ is closed under complementation and countable union. μ is a function from $\mathscr{F}$ to [0,1] such that $\mu(S) = 1$ and $\mu(\alpha \cup \beta) = \mu(\alpha) + \mu(\beta)$, for disjoint α and β. We then let $[\![(\phi \geqslant \psi)]\!]^w_{\mathscr{M}} = 1$ iff $\mu([\![\phi]\!]_{\mathscr{M},S}) \geq \mu([\![\psi]\!]_{\mathscr{M},S})$.

Probabilistic semantics avoids the disjunction puzzle and also validates a range of intuitively valid inference patterns identified by Yalcin (2010).[12] As a result, there appear to be solid grounds for favoring probabilistic over world-ordering semantics.[13]

5.3 HOLLIDAY AND ICARD'S ALTERNATIVE

As Kratzer (2012) notes, her (1991*b*) choice of lifting operation is one among many. It thus remains to be seen whether one can formulate an alternative lifting operation that yields better predictions.

[11] This example is due to Yalcin (2010).

[12] See Lassiter (2010, 2011, 2015) and Yalcin (2010).

[13] My focus in this paper is the debate between probabilistic and world-ordering semantics, so I will leave aside the question of whether a probabilistic semantics should be based on finite additivity, countable additivity, or qualitative additivity (see Holliday and Icard (2013) and Lassiter (2015) for discussion of qualitative additivity; see §5.5 for discussion of countable additivity). I also leave aside the question of whether probabilistic semantics fares better than what Holliday and Icard (2013) call *event-ordering semantics*.

Holliday & Icard (2013) claim to do just that: they present an alternative lifting operation that promises to resolve the disjunction puzzle and capture the core inferences involving probability talk:

> m-lifting: $[\![\phi]\!]_{\mathscr{M},S} \succeq^m [\![\psi]\!]_{\mathscr{M},S}$ iff there exists an injective function $f : [\![\psi]\!]_{\mathscr{M},S} \to [\![\phi]\!]_{\mathscr{M},S}$ such that $\forall w \in [\![\psi]\!]_{\mathscr{M},S} : f(w) \succeq w$.

An injective function is one that maps every distinct element in the domain to a distinct element in the codomain: that is, for every w, w' in the domain, if $f(w) = f(w')$, then $w = w'$. Thus, m-lifting tells us that a proposition α is at least as highly ranked as β iff each β-world can be mapped to a distinct α-world that is at least as highly ranked.

It's easy to see how this new lifting operation resolves the disjunction puzzle. Consider an instance of I1: $((p \gtrdot q) \wedge (p \gtrdot \neg q)) \to (p \gtrdot (q \vee \neg q))$. Suppose $W = S = \{w, w'\}$, where w is the sole p-world and q-world. Our instance of I1 will be false at w if $w \succeq w'$.

Intuitively, m-lifting invalidates I1 because the same p-world can't do double duty in matching up to each of the q and $\neg q$-worlds: this is ruled out by the requirement that each $(q \vee \neg q)$-world be mapped to a *distinct* p-world in order for p to be at least as likely as the disjunction.

Moreover, Holliday & Icard claim that world-ordering semantics with m-lifting validates all of the core, valid inferences on Yalcin's (2010) list. Consequently, semanticists will have to look elsewhere for grounds favoring probabilistic over world-ordering semantics.

To be clear: the threat here is not that there are *no* grounds for favoring one semantics over the other. There are other motivations for probabilistic semantics, and there are questions about whether m-lifting can capture other data about the use of probability modals.[14] But the general lesson of Holliday & Icard (2013) is thought to be the following: whatever the grounds will be for favoring one semantics over the other, they will not concern the logic of comparative likelihood.

[14] See Lassiter (2010, 2011, 2015), Moss (2013, 2015), Rothschild (2012), Swanson (2006, 2011, 2016), and Yalcin (2007, 2011) for discussion of alternative motivations for probabilistic semantics. See Lassiter (2015) for criticisms of m-lifting distinct from those I raise below. Note also that caution must be applied when integrating m-lifting into Kratzer's larger account of modal language. I1 is arguably valid for deontic comparatives, so m-lifting cannot serve as a general lifting operation for all comparative modal language. Thanks to Eric Swanson for noting this point.

5.4 THE COUNTERMODEL

As it turns out, the dialectic is more complex. Holliday & Icard's semantics still fails to validate one intuitively valid inference pattern. The inference pattern in question is what Yalcin (2010) calls V11:

V11:
P1. $\phi \geqslant \psi$
P2. $\Delta\psi$
C. $\Delta\phi$

V11 is clearly a valid inference pattern. For example, if rain is at least as likely as high winds, then if high winds are likely, rain is likely as well.

But the following constitutes a countermodel to V11 for world-ordering semantics with m-lifting:

Let $\mathscr{M} = \langle W, S, V, \succeq, \uparrow \rangle$, where:
$W = S = \mathbb{N}$
$V(p) = \{x \in \mathbb{N} : x \text{ is even}\}$
$\succeq$ is a flat ranking: for every $w, w' \in S, w \succeq w'$
$\uparrow = m$-lifting

Choose any $w \in W$. The following instance of V11 is false at w:

(*) $((p \geqslant (p \vee \neg p)) \wedge \Delta(p \vee \neg p)) \rightarrow \Delta p$

To see why, notice that since every pair of worlds in S is equally ranked, establishing the first conjunct of the antecedent of (*) simply requires showing that there exists an injection from the $(p \vee \neg p)$-worlds to the p-worlds. Here is such an injection: $f(x) = 2x$. Next, note that the second conjunct of the antecedent is true at w since there is trivially an injection from the empty set to $\mathbb{N}$, but not vice versa. However, the consequent of the conditional is

false at w since $(\neg p \geqslant p)$ is true at w: $f(x) = (x + 1)$ is an injection from the p-worlds to the $\neg p$-worlds.[15]

What drives this countermodel is the following distinctive feature of infinitely large sets: a proper subset of a set may have the same cardinality as the set itself. In this case, $V(p)$ is a subset of $\mathbb{N}$ that has the same cardinality as $\mathbb{N}$. Hence, since each element of $\mathbb{N}$ is ranked equally high, p will be at least as likely as $\mathbb{N}$. But the entire domain, $\mathbb{N}$, is trivially likely, while p is not, since its complement is equally likely. Hence, the semantics fails to validate V11.[16]

In contrast, it's straightforward to verify that probabilistic semantics validates V11 regardless of the size of S:

Proof. If $[\![\Delta\psi]\!]^w_{\mathscr{M}} = 1$, then $\mu([\![\psi]\!]_{\mathscr{M},S}) > .5$. Hence, if $[\![(\phi \geqslant \psi)]\!]^w_{\mathscr{M}} = 1$, then $\mu([\![\phi]\!]_{\mathscr{M},S}) \geq \mu([\![\psi]\!]_{\mathscr{M},S})$, in which case $\mu([\![\phi]\!]_{\mathscr{M},S}) > .5$, and thus $[\![\Delta\phi]\!]^w_{\mathscr{M}} = 1$.

It may be helpful to re-state these points with the aid of a simple interpretation of our countermodel. Let every $x \in \mathbb{N}$ stand for a world containing x stars.[17] p therefore expresses the proposition that the number of stars is even. Our countermodel then demonstrates that world-ordering semantics with m-lifting invalidates the following inference when S is infinitely large:

P1. An even number of stars is at least as likely as an even or odd number of stars.

P2. An even or odd number of stars is likely.

C. An even number of stars is likely.

Probabilistic semantics validates the inference nonetheless. Thus, probabilistic semantics has a clear advantage over even the most sophisticated, world-ordering alternative.

[15] Our model also constitutes a countermodel to what Holliday & Icard call V13:

V13:
P1. $(\phi \wedge \neg\psi) > \bot$
C. $(\phi \vee \psi) > \psi$

The following instance is false at any $w \in W$: $(((p \vee \neg p) \wedge \neg p) > \bot) \rightarrow (((p \vee \neg p) \vee p) > p)$. See §5.6 for further discussion of V13.

[16] We can construct a similar countermodel with an uncountably infinite domain. Let $W = S = [0, 2\pi)$, let $V(p) = [0, \pi]$, and retain the rest of our original countermodel. (*) is false in this model for the same reason: the cardinality of $V(p)$ is the same as that of S and that of $[\![\neg p]\!]$.

[17] This interpretation is inspired by an example from Portner (2009, 33).

5.5 POSSIBLE REPLIES

As we saw in the previous section, the shift to m-lifting does not suffice to capture all of the intuitively valid inferences involving probability talk. However, it's worth exploring whether defenders of world-ordering semantics can find a way to undermine the significance of our countermodel. I consider three replies below.

5.5.1 Add a Density Measure

Implicit in m-lifting is the thought that judgments of comparative likelihood involve comparing the number of worlds in which each proposition is true. Requiring there to exist a certain injection from $[\![\psi]\!]_{\mathscr{M},S}$ to $[\![\phi]\!]_{\mathscr{M},S}$ in order to verify $(\phi \geqslant \psi)$ is one way of capturing this thought: there exists such a function only if the cardinality of $[\![\phi]\!]_{\mathscr{M},S}$ is at least as great as that of $[\![\psi]\!]_{\mathscr{M},S}$. But one might take our countermodel to illustrate that cardinality is too crude a notion to play this role: intuitively, there are fewer evens than natural numbers, but their cardinalities are the same.

One might instead appeal to the notion of *density* to capture this intuition: any given stretch of natural numbers will contain fewer evens than natural numbers. Perhaps such judgments of density underlie our judgments of comparative likelihood: α's being at least as likely as β requires not just that there be an injection from the β-worlds to the α-worlds, but that the density of the α-worlds be at least as great as the density of the β-worlds. Such a lifting operation will block our countermodel, since the density of the natural numbers is greater than that of the evens.

But the problem with this reply is that any rigorous specification of such a density measure will invoke precisely the sort of mathematically sophisticated tacit knowledge that defenders of non-probabilistic semantics seek to avoid (e.g. tacit knowledge of limits, and so on). Moreover, the very idea of density is that of a frequency within a reference class, so any appeal to tacit knowledge of density presupposes a tacit grasp of quantitative probability.

5.5.2 Infinite Domains are Irrelevant

Our countermodel crucially relies on the distinctive characteristics of infinitely large sets in order to verify the antecedent of (*): the cardinality of a proper subset can be equal to the cardinality of the entire set itself. And indeed, Holliday & Icard's semantics validates V11 over the class of models in which S is finite. Perhaps such a restriction is warranted, given the aims of natural language semantics: infinite domains play no role in the thoughts of the layman.

But it's hard to give this reply much credence. To be sure, it's common in semantics to focus on models with finite domains: for instance, semanticists will often stipulate a finite domain of epistemically accessible worlds in order to secure the Limit Assumption and so simplify their semantics for deontic modals.[18] But such restrictions are mere stipulations—and for good reason.

Consider the aforementioned interpretation of our countermodel: one can wonder about the number of stars in the universe, and as long as one is not willing to place a finite upper bound on the answer, one's information state is best modeled by a countably infinite domain of possible worlds, each containing an ever-greater number of stars.[19] Portner (2009, 32–33) goes so far as to claim that the domain of accessible worlds recognized by ordinary language users is *typically* infinite. But it suffices for our purposes to note that the domain may *sometimes* be infinite, and thus it would be a mistake to declare our countermodel irrelevant on the grounds that ordinary speakers never traffic in infinite domains.

5.5.3 Models are Path-Finite

In a post-print of their original (2013), Holliday & Icard add the following assumption:

> [W]e assume world-ordering models are *path-finite*, i.e., there is no infinite path $w_1 \preceq w_2 \preceq w_3 \ldots$ of distinct worlds, just as with a finitely additive measure, there is no infinite

[18] Note that our countermodel is consistent with the Limit Assumption.
[19] We can stipulate that one knows or believes that the number of stars is finite.

sequence of distinct worlds with non-zero, non-decreasing measure (post-print, 526).[20]

This assumption allows for infinite domains but will still block our countermodel: a model with a flat ranking on an infinite domain is not path-finite. And indeed, V11 is valid in the class of world-ordering models with path-finite rankings.

However, this strategy has several costs. First, the path-finiteness requirement is *ad hoc*: it is motivated only on the grounds that it validates an inference that the semantics would otherwise render invalid. Now, Holliday & Icard do draw an analogy between the path-finiteness requirement and a similar constraint on probabilistic semantics: "with a finitely additive measure, there is no infinite sequence of distinct worlds with non-zero, non-decreasing measure" (post-print, 526). But this analogy is flawed. One cannot have a finitely additive, non-zero, non-decreasing measure over an infinite domain—but one *can* have a finitely additive measure over an infinite domain that assigns probability zero to each outcome. Indeed, one of the principal motivations for finite but not countable additivity is that it allows one to assign probability zero to each outcome in a countably infinite domain and thereby capture the judgment that each outcome is equally likely.[21] Furthermore, if the sample space is continuous—i.e. if it contains an uncountably infinite set of possible outcomes—then we *must* assign probability zero to an uncountable number of outcomes.[22]

Defenders of probabilistic semantics are aware of these facts. Several of them explicitly allow for assigning probability zero to every outcome in an infinite domain.[23] Thus, without further argument, there is no reason to believe that defenders of probabilistic semantics have to make a stipulation analogous to the path-finiteness requirement. The former theorists can allow for assigning probability zero to every outcome in an infinite domain, but Holliday & Icard cannot allow for a flat ranking over an infinite domain: as we saw in §5.4, models with such rankings invalidate V11.

The second problem with the path-finiteness requirement concerns the origin of the ranking on worlds. The standard account—due to Kratzer (1981; 1991; 2012)—is that the

[20]Post-print available at `https://philosophy.berkeley.edu/file/869/Holliday_and_Icard_SALT_23.pdf` (last accessed Feb. 16, 2017). See Harrison-Trainor, Holliday and Icard (forthcoming) and Harrison-Trainor, Holliday and Icard (2017) for further development of path-finite, world-ordering semantics.
[21]See de Finetti (1974).
[22]See Williamson (2007).
[23]See Cariani (2016), Carr (2015), and Yalcin (2007).

ranking on worlds is fixed by a contextually-determined set of ordering source propositions, as we reviewed in §5.2. But suppose the set of ordering source propositions is empty: every world will trivially verify the same ordering source propositions and will thus be equally ranked. Consequently, if the domain of epistemic possibilities is infinite, the path-finiteness requirement precludes an empty ordering source (recall that a flat ranking on an infinite domain is not path-finite). However, this is a bad result. Possibility modals are often thought to have readings that involve an empty ordering source, so why should the situation be any different for probability modals, or different when the domain is infinite? The ordering source for probability modals is supposed to represent a contextually determined standard of normality. But the context of use might simply fail to determine some such standard and thereby leave the ordering source empty. The path-finiteness requirement thus conflicts with the standard account of what determines the ranking on worlds in the first place: the mere fact that the domain is infinite should not rule out the possibility of an empty ordering source.

The final problem with the path-finiteness requirement concerns our intuitions about examples involving infinite domains. Recall the star interpretation of our countermodel discussed above: suppose we are wondering how many stars there are in the universe, and we are unwilling to place a finite upper bound on the answer. Here our information state is best modelled by a countably infinite domain of possible worlds, each containing an ever-greater number of stars. Now, *prima facie*, it is possible to be in a coherent information state of this sort according to which, above a certain threshold, any number of stars is as likely as any other number of stars. Or suppose we are wondering about the precise value of some physical constant. Again, it is plausible that there exists some state of belief or evidence according to which each of an infinite range of values is equally likely. One candidate for such an information state is that of total ignorance.[24] There are others. One might possess—

[24]It is controversial whether a single probability measure can adequately represent a state of total ignorance. A referee notes that there is no constructive proof of a uniform distribution over the natural numbers (see Lauwers (2009)). And the non-constructive choices required to generate such a distribution seem at odds with the distribution's modelling a state of total ignorance. But as the referee also notes, modelling ignorance by a set of probability measures may help avoid this problem. See Rothschild (2012) for an example of a probabilistic semantics that relies on sets of measures. Finally, it may be easier to model ignorance when the domain is uncountably infinite—e.g. consider the Lebesgue measure on [0,1], on which the probability of every interval in [0,1] is equal to its length.

or merely *believe* that one possesses—positive evidence that the value of this constant is determined by a random process. All of this is to say: our semantics for probability modals should not rule out the possibility that each of an infinite set of outcomes is equally likely.

But the path-finiteness requirement does rule this out. Equiprobability of outcomes requires a flat ranking—one that is not path-finite if the domain of epistemic possibilities is infinitely large.[25] By contrast, probabilistic semantics allows for equiprobability of outcomes across an infinite domain. One need only assign each outcome probability zero.[26]

To sum up: the path-finiteness requirement secures V11 but at the expense of (a) introducing an *ad hoc* constraint on world-orders; (b) precluding an empty ordering source when the domain is infinite; (c) ruling out the possibility of equiprobable outcomes across an infinite domain.

5.6 PARTNERS IN CRIME?

In defending probabilistic semantics, I've twice appealed to probability-zero epistemic possibilities. But one might wonder whether allowing such possibilities leads to undesirable consequences. If so, one could argue that a viable probabilistic semantics must indeed make a stipulation analogous to the path-finiteness requirement: models cannot include a measure μ such that $\mu(\{w_1\}) \leq \mu(\{w_2\}) \leq \mu(\{w_3\}) \ldots$ for an infinite sequence of distinct worlds. Such a measure is possible only if each world is assigned probability zero (any greater value would violate the requirement that $\mu(S) = 1$).

So: what sort of undesirable consequences might follow from allowing probability-zero possibilities? It is true that if we allow for probability-zero possibities, leaving the rest of our probabilistic semantics unchanged, our theory will invalidate a plausible principle connecting epistemic possibility and comparative likelihood:

[25] A world-order in which each possibility is incomparable will deliver the result that none of the outcomes are more or less likely than the others. However, incomparability of outcomes is not the same as equiprobability—i.e. that each is equally likely.

[26] Equiprobability of outcomes across a countably infinite domain requires a finitely but not countably additive measure. Equiprobability across an uncountably infinite domain is consistent with countable additivity.

Regularity$_C$: $\Diamond\phi \rightarrow (\phi > \bot)$[27]

Recall that on the probabilistic semantics discussed in §5.2, facts about comparative likelihood are settled by comparing the measure value of each proposition. Thus, since contradictions receive probability zero, no probability-zero possibility will be more likely than a contradiction. This is a highly unintuitive result. Returning to our star case above, it seems absurd to claim that a googol stars in the universe is no more likely than $0 = 1$.[28]

However, it is possible to amend our probabilistic semantics to validate *Regularity*$_C$ while allowing for probability-zero possibilities. To see how, first distinguish *Regularity*$_C$ from a similar principle, also called "Regularity", that is often discussed in formal epistemology and probability theory:

Regularity$_P$: if $\alpha \neq \emptyset$ and $\alpha \in \mathscr{F}$, then $\mu(\alpha) > 0$.[29]

Regularity$_P$ simply expresses a constraint on probability measures, yet it's natural to think that *Regularity*$_P$ is the only way to secure *Regularity*$_C$ in a probabilistic semantics. If so, one cannot validate *Regularity*$_C$ if one accepts probability-zero possibilities: such possibilities violate *Regularity*$_P$, since they would be non-empty members of $\mathscr{F}$ that do not receive greater-than-zero probability.

But the natural thought is false. There are other routes to validating *Regularity*$_C$ in a probabilistic semantics. Consider what I'll call the *Modified Probabilistic Semantics*:

$$\begin{aligned}[\![(\phi \geqslant \psi)]\!]^w_{\mathscr{M}} = 1 \quad \text{iff} \quad & \text{(i) } \mu([\![\phi]\!]_{\mathscr{M},S}) \geq \mu([\![\psi]\!]_{\mathscr{M},S}), \\ & \text{(ii) } [\![(\Diamond\psi \rightarrow \Diamond\phi)]\!]^w_{\mathscr{M}} = 1, \text{ and} \\ & \text{(iii) } [\![(\Box\psi \rightarrow \Box\phi)]\!]^w_{\mathscr{M}} = 1.\end{aligned}$$

This semantics ensures that a contradiction is never at least as likely as an epistemically possible proposition—even if this proposition has probability zero. But any proposition is

[27] The subscript indicates that the principle constrains judgments of comparative likelihood; I discuss a different type of regularity principle below.

[28] This case is adapted from an example due to Wesley Holliday.

[29] "Regularity" is sometimes formulated as the stronger thesis that all non-empty subsets of S receive positive probability. See Hájek (ms.) for discussion.

at least as likely as a contradiction, so an epistemically possible proposition will always be more likely than a contradiction.[30] The semantics also ensures that epistemically necessary propositions will be more likely than those that merely receive probability one. This is a desirable result: additivity requires that the negation of any probability-zero possibility receive probability one, but such propositions—e.g. the proposition that the number of stars is not equal to a googol—are clearly less likely than epistemic necessities, such as the proposition that the number of stars is either even or odd. Finally, the semantics validates V11 and the other validities on Yalcin's (2010) list.

Now, the semantics will fail to validate what Holliday & Icard call V13:

V13:
P1. $(\phi \wedge \neg\psi) > \bot$
C. $(\phi \vee \psi) > \psi$

But it's not obvious that this inference is actually valid. I take it that V13 is not self-evident in the way that V11 is. Rather, V13 reflects something like the following line of reasoning: if $(\phi \wedge \neg\psi)$ is more likely than a contradiction, then it's possible for ϕ to occur without ψ; but then there are more ways for $(\phi \vee \psi)$ to be true than for ψ itself to be true—namely, all of the ψ-ways plus the $(\phi \wedge \neg\psi)$-ways. This is a valid line of reasoning when the domain of possibilities is finite. Not so when the domain is infinite. Suppose ϕ expresses the proposition that the number of stars is a multiple of 3, and ψ expresses the proposition that the number of stars is a multiple of 6. It is possible for the number of stars to be a multiple of 3 and not a multiple of 6, but the cardinality of $\{w$: the number of stars in w is a multiple of 3 or a multiple of 6$\}$ is the same as the cardinality of $\{w$: the number of stars in w is a multiple of 6$\}$. Thus, there is a perfectly respectable notion of size according to which the above line of reasoning goes wrong when the domain is infinite. Of course, others might wish to understand size in terms of the proper superset relation, in which case there is nothing wrong with the above line of reasoning even in the case of an infinite domain. But this only shows that we reach a standoff over V13.

[30] See Easwaran (2014) for alternative ways of capturing *Regularity*$_C$ without *Regularity*$_P$.

Still, it might be thought that our semantics falters on a related point. Surely there being 10^{10} *or* 10^{11} stars is more likely than there being 10^{11} stars, but if each disjunct is assigned probability zero, the disjunction will be just as likely as either disjunct.[31] Again though, infinite domains reveal a problem. Intuitively, the disjunction is more likely because there are more ways for it to be true. But it's easy to miss that when the domain is infinite, the disjunction and each disjunct are still *false* in the same number of worlds—namely, a countably infinite number. We can then reason as follows. The disjunction and each disjunct are just as likely to be false, since they are false in the same number of worlds and there are no grounds for thinking that the set of worlds in which one is false is more likely than the set of worlds in which the other is false. And since the disjunction and each disjunct are equally likely to be false, they are equally likely to be true.

What this shows is that we essentially confront a puzzle. Each of the following is plausible, but they are jointly inconsistent:

- The disjunction is more likely than either disjunct.
- The disjunction and each disjunct are equally likely to be false.
- If two sentences are equally likely to be false, then they are equally likely to be true.

I suggest we reject the first. Its plausibility results from the failure to recognize that when the domain is infinite, each sentence is false in the same number of worlds.

It remains to show that our semantics is not *ad hoc*: it is motivated by considerations other than those of securing $\textit{Regularity}_C$ while allowing for probability-zero possibilities. The general idea behind this semantics is that judgments of comparative likelihood are sensitive to the epistemic possibility or necessity of the propositions so compared. Much of the significance of probability talk stems from its use as a guide in the formation of our attitudes towards non-modalized propositions: the question of whether α is at least as likely as β is ultimately aimed at addressing the questions of whether α and whether β. So it would be surprising if judgments of comparative likelihood essentially throw out the information about whether α or β is epistemically possible, or epistemically necessary—information that is clearly relevant to the question of whether α and whether β. But this is exactly what

[31] cf. Pruss (2014).

would happen if we simply compare the measure values of each proposition: each might have probability zero while only one is epistemically possible, and each might have probability one while only one is epistemically necessary. Similarly, consider how we often go about determining whether α is at least as likely as β: we check to see whether the grounds favoring α are at least as strong as those favoring β. If this is how we make judgments of comparative likelihood, it is no surprise that the epistemic possibility or necessity of each proposition bears on such judgments. For if either is epistemically *im*possible, this fact undercuts any evidence in its favor. And if either is epistemically necessary, the other must also be epistemically necessary if it is to be supported by grounds of equal strength.

I take the upshot to be the following. Defenders of world-ordering semantics cannot find partners in crime to mitigate the damages of requiring path-finite world-orders: there exists a well-motivated probabilistic semantics that validates V11, allows for equiprobable outcomes across an infinite domain, and retains $\mathit{Regularity}_C$. As a result, probabilistic semantics still provides the best account of the inference patterns governing natural language probability talk.

To be sure, nothing I've said rules out the possibility that an alternative lifting operation will give us everything we want. But it remains an open question whether such an operation exists. For now, then, the data favors a probabilistic semantics.

APPENDIX A

NON-STANDARD RELATIVISM

In §3.4.3, I showed how cloudy contextualism provides a unified explanation of the shared eavesdropping data. Here I show that one can also give a unified explanation of the data in a relativist framework. We saw in §3.3.2 that the **Heterodox Semantics** is not the right semantics in which to develop relativism about reason claims, since this semantics makes the wrong predictions about how reason claims embed. But there is another option, one that makes use of what Kratzer (1977, 1981, 2012) calls a "modal base"—i.e. a function f from worlds to (simplifying a bit) bodies of information (i.e. sets of worlds).

Suppose we use a modal base to give a parallel compositional semantics for reason and modal claims:

> *might* ϕ is true at a context, modal base, and world triple $<c,f,w>$ iff $f(w)$ is compatible with the proposition expressed by ϕ at c.
>
> *there is reason to believe that* ϕ is true at $<c,f,w>$ iff $f(w)$ counts in favor of believing the proposition expressed by ϕ at c.[1]

We can then account for the difference in embedding data by positing different default constraints on the value of f: for reason claims, the value of this parameter will be a function tracking the information contained in epistemic states; for modal claims, the value of this

[1]This semantics for reason claims is just a sketch of the general shape such a theory should take. Some might wish to take the counting-in-favor-of relation as primitive; others might wish to reduce it to a probabilistic relation. There is also a question about whether reason claims should have a quantifier in their semantics, or whether their surface grammar is misleading. I take up these issues in Chapter 2.

parameter will be a function tracking relevant truths, where there is no requirement that such truths be contained in epistemic states. On this semantics, reason and modal claims emerge as two different "flavors" of modal language in general.

To see how this semantics works, consider the following sentences:

(1) John might be infected.

(2) There is reason to believe that John is infected.

The modal base for (1) might be a function from a world w to the set of worlds consistent with John's symptoms at w. The modal base for (2) might be a function from a world w to the set of worlds consistent with what the speaker of the context knows at w. The reason claim will then be true at a world depending on how things stand with what the speaker knows at that world, while the modal claim will be true at a world depending on how things stand with John's symptoms at that world. These propositions will likely be non-equivalent. Consider a world at which John's symptoms rule out the presence of an infection but the speaker possesses misleading evidence suggesting that John has the characteristic symptoms of the infection. This will be a world at which (1) is false but (2) is true.

Now, to give a compositional semantics is not yet to settle what proposition (if any) is expressed by a typical use of a sentence.[2] One might therefore employ our above semantics in the context of two very different theories of the objects of assertion. One such theory integrates our semantics with a contextualist account of the information conveyed by reason and modal claims. On this account, the context of use settles the value of the modal base parameter f, and a use of a reason or modal claim will express an ordinary proposition (i.e. a set of worlds). However, one might instead take the value of f to be settled by the context of assessment, not the context of use. On this account, a use of a reason or modal claim will not express an ordinary proposition. Instead, the use will express a formal object that is modal base-neutral and so will be true relative to some modal base-world pairs and false relative to others. The result is a kind of non-standard relativism—non-standard in the sense that current relativist theories are developed in the framework of the **Heterodox Semantics**. This type of relativist theory will then explain the eavesdropping data in a

[2]This point is emphasized in Dummett (1978, Chapter 1), Lewis (1980), Ninan (2010), and Yalcin (2014).

unified and familiar way: the speaker is warranted in making a reason or modal claim if this claim true at the context of assessment she presently occupies, but the eavesdropper is warranted in rejecting the speaker's assertion if the claim is false at the eavesdropper's context of assessment.

This type of non-standard relativism opens up the possibility of a different kind of **Split View** than the one we discussed in §3.4.1. Suppose one adopts a relativist implementation of the **Heterodox Semantics** for modal claims and our non-standard relativism for reason claims. The resulting view will offer a unified, relativist explanation of the eavesdropping data. This view can also explain the difference in embedding data: the view can employ the existing **Heterodox Semantics** explanation of the unembeddability of epistemic contradictions and can co-opt the explanation of the embeddability of reason contradictions I gave above. Now, I have no knock-down objection to this proposal, but let me explain why I find it less satisfying than the proposal I've given in the main text. The central drawback to this type of **Split View** is its lack of theoretical economy. The **Heterodox Semantics** for modal claims is standardly motivated by its ability to explain the eavesdropping and epistemic contradiction data about modal claims—data that is alleged to be difficult to accommodate within a non-heterodox semantics for modal claims (see, for instance, MacFarlane (2011, 2014) and Yalcin (2007, 2011)). However, I've just argued that one can explain each piece of data in the context of the standard, Kratzerian compositional semantics for modal claims. It is far simpler, then, to just give a single compositional semantics for both reason and modal claims, as I've done above.

APPENDIX B

FORMAL PROOFS

Here I present a formal argument for why any plausible extension of the **Heterodox Semantics** will predict that reason contradictions are incoherent to entertain.

Let $R\phi$ be an abbreviation of *there is reason to believe that* ϕ, let *No-*$R\phi$ be an abbreviation of *there is no reason to believe that* ϕ, let c be a variable ranging over contexts, let w range over metaphysical possibilities, let s range over information states—i.e. sets of metaphysical possibilities—and let $[\![\cdot]\!]$ be the interpretation function—i.e. a function from expressions to their semantic values.[1]

The following theses jointly entail that no coherent supposition state "accepts" a reason contradiction, where a supposition state s accepts that ϕ iff $\forall w \in s : [\![\phi]\!]^{c,s,w}$ is true:

(a) For all points of evaluation $<c,s,w>$, if $[\![R\phi]\!]^{c,s,w}$ is true, then $[\![\textit{No-}R\phi]\!]^{c,s,w}$ is false.

(b) For all $<c,s,w>$, if $s \subseteq [\![\phi]\!]^{c,s}$, then $[\![R\phi]\!]^{c,s,w}$ is true.

Proof. Suppose s accepts that ϕ. By definition, s accepts that ϕ iff $\forall w \in s : [\![\phi]\!]^{c,s,w}$ is true iff $\forall w \in s : w \in [\![\phi]\!]^{c,s}$ iff $s \subseteq [\![\phi]\!]^{c,s}$. Hence, by (b), $[\![R\phi]\!]^{c,s,w}$ is true, for all worlds w. Whence, by (a), $[\![\textit{No-}R\phi]\!]^{c,s,w}$ is false, for all worlds w. Thus, $\forall w \in s : [\![\textit{No-}R\phi]\!]^{c,s,w}$ is false. But by definition, s accepts that *No-*$R\phi$ only if $\forall w \in s : [\![\textit{No-}R\phi]\!]^{c,s,w}$ is true. Hence,

[1] Here is an example of how the interpretation function works. Take an expression ϕ. In standard two-dimensional semantics, expressions are assigned extensions (e.g. truth-values) relative to a context c and, in a simple case, an index consisting of a world parameter w. $[\![\phi]\!]^c$ would then be a function from worlds to truth-values, or equivalently, a set of possible worlds—namely, those at which the proposition expressed by ϕ at c is true.

s accepts that $\phi \wedge \textit{No-R}\phi$ only if $\forall w \in s : [\![\textit{No-R}\phi]\!]^{c,s,w}$ is false and $\forall w \in s : [\![\textit{No-R}\phi]\!]^{c,s,w}$ is true. So s accepts that $\phi \wedge \textit{No-R}\phi$ only if $s = \emptyset$. But s was arbitrary, so for all s, s accepts that $\phi \wedge \textit{No-R}\phi$ only if $s = \emptyset$. $\square$

See §3.3.2 for my arguments that any plausible, heterodox-style semantics for reason claims must accept (a) and (b).

BIBLIOGRAPHY

Alvarez, Maria. 2010. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford University Press.

Anderson, Charity. 2014. "Fallibilism and the Flexibility of Epistemic Modals." *Philosophical Studies* 167(3):597–606.

Cariani, Fabrizio. 2016. Deontic Modals and Probability: One Theory to Rule Them All? In *Deontic Modality*, ed. Nate Charlow and Matthew Chrisman. Oxford University Press.

Cariani, Fabrizio, Magdalena Kaufmann and Stefan Kaufmann. 2013. "Deliberative Modality Under Epistemic Uncertainty." *Linguistics and Philosophy* 36(3):225–259.

Carnap, Rudolf. 1947. *Meaning and Necessity*. University of Chicago Press.

Carr, Jennifer. 2015. "Subjective Ought." *Ergo, an Open Access Journal of Philosophy* 2.

Cohen, Stewart. 1988. "How to Be a Fallibilist." *Philosophical Perspectives* 2(n/a):91–123.

Comesaña, Juan and Matthew McGrath. 2014. Having False Reasons. In *Epistemic Norms: New Essays on Action, Belief, and Assertion*, ed. Clayton Littlejohn and John Turri. Oxford University Press.

Crabill, Joshua D. 2013. "Suppose Yalcin is Wrong About Epistemic Modals." *Philosophical Studies* 162(3):625–635.

de Finetti, Bruno. 1974. *Theory of Probability, Vols 1 and 2*. New York: John Wiley.

DeRose, K. 1998. "Simple 'Might's, Indicative Possibilities and the Open Future." *Philosophical Quarterly* 48(190):67–82.

DeRose, Keith. 1991. "Epistemic Possibility." *Philosophical Review* 100:581–605.

DeRose, Keith. 2017. Contextualism and Fallibilism. In *The Routledge Handbook of Epistemic Contextualism*, ed. J. Ichikawa. Routledge.

Dodd, Dylan. 2010. "Confusion About Concessive Knowledge Attributions." *Synthese* 172(3):381–396.

Dodd, Dylan. 2011. "Against Fallibilism." *Australasian Journal of Philosophy* 89(4):665–685.

Dorr, Cian and John Hawthorne. 2013. "Embedding Epistemic Modals." *Mind* 122(488):867–913.

Dougherty, Trent. 2011. Fallibilism. In *The Routledge Companion to Epistemology*, ed. Duncan Pritchard and Sven Bernecker. Routledge.

Dougherty, Trent and Patrick Rysiew. 2009. "Fallibilism, Epistemic Possibility, and Concessive Knowledge Attributions." *Philosophy and Phenomenological Research* 78(1):123–132.

Dougherty, Trent and Patrick Rysiew. 2011. "Clarity About Concessive Knowledge Attributions: Reply to Dodd." *Synthese* 181(3):395–403.

Dowell, J. L. 2011. "A Flexible Contextualist Account of Epistemic Modals." *Philosophers' Imprint* 11(14):1–25.

Dowell, J.L. ms. "Contextualism about Epistemic Modals." Unpublished manuscript.

Dummett, Michael A. E. 1978. *Truth and Other Enigmas*. Cambridge, Mass.: Harvard University Press.

Easwaran, Kenny. 2014. "Regularity and Hyperreal Credences." *Philosophical Review* 123(1):1–41.

Egan, Andy. 2007. "Epistemic Modals, Relativism and Assertion." *Philosophical Studies* 133(1):1–22.

Egan, Andy, John Hawthorne and Brian Weatherson. 2005. Epistemic Modals in Context. In *Contextualism in Philosophy*, ed. G. Preyer and G. Peter. Oxford University Press pp. 131–170.

Fantl, Jeremy and Matthew McGrath. 2009. *Knowledge in an Uncertain World*. Oxford University Press.

Feldman, Richard. 1981. "Fallibilism and Knowing That One Knows." *Philosophical Review* 90(2):266–282.

Finlay, Stephen. 2014. *Confusion of Tongues: A Theory of Normative Language*. Oxford: Oxford University Press.

Greco, Daniel. 2014. "A Puzzle About Epistemic Akrasia." *Philosophical Studies* 167(2):201–219.

Hacking, Ian. 1967. "Possibility." *Philosophical Review* 76(2):143–168.

Hájek, Alan. ms. "Staying Regular?" Unpublished manuscript.

Hamblin, C. L. 1959. "The Modal "Probably"." *Mind* 68(270):234–240.

Harrison-Trainor, Matthew, Wesley Holliday and Thomas Icard. 2017. Preferential Structures for Comparative Probabilistic Reasoning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*.

Harrison-Trainor, Matthew, Wesley Holliday and Thomas Icard. forthcoming. "Inferring Probability Comparisons." *Mathematical Social Sciences* .

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford University Press.

Hawthorne, John. 2007. "Eavesdroppers and Epistemic Modals." *Philosophical Issues* 17(1):92–101.

Hawthorne, John. 2012. "Knowledge and Epistemic Necessity." *Philosophical Studies* 158(3):493–501.

Hawthorne, John and Jason Stanley. 2008. "Knowledge and Action." *The Journal of Philosophy* 105(10):571–590.

Henning, Tim. 2014. "Normative Reasons Contextualism." *Philosophy and Phenomenological Research* 88(3):593–624.

Holliday, Wesley H. and Thomas F. Icard. 2013. "Measure Semantics and Qualitative Semantics for Epistemic Modals." *Proceedings of SALT 23* pp. 514–534.

Hornsby, Jennifer. 2008. A Disjunctive Conception of Acting for Reasons. In *Disjunctivism: Perception, Action, Knowledge*, ed. Adrian Haddock and Fiona Macpherson. Oxford University Press.

Ichikawa, Jonathan. 2011. "Quantifiers and Epistemic Contextualism." *Philosophical Studies* 155(3):383–398.

Khoo, Justin. 2015. "Modal Disagreements." *Inquiry* (5):1–24.

Kolodny, Niko. ms. "Lectures on Reasons and Rationality." Unpublished manuscript.

Kolodny, Niko and John MacFarlane. 2010. "Ifs and Oughts." *Journal of Philosophy* 107(3):115–143.

Koopman, Bernard O. 1940. "The Axioms and Algebra of Intuitive Probability." *Annals of Mathematics* pp. 269–292.

Kotzen, Matt. ms. "Conditional Oughts and Simpson's Paradox." Unpublished manuscript.

Kratzer, Angelika. 1977. "What *Must* and *Can* Must and Can Mean." *Linguistics and Philosophy* 1(3):337–55.

Kratzer, Angelika. 1981. The Notional Category of Modality. In *Words, Worlds, and Contexts, new approaches to word semantics*, ed. Eikmeyer and Rieser. Berlin: Walter de Gruyter pp. 38–74.

Kratzer, Angelika. 1991*a*. Conditionals. In *Handbuch Semantik/Handbook Semantics*, ed. Stechow and Wunderlich. Walter de Gruyter.

Kratzer, Angelika. 1991*b*. Modality. In *Semantics: An International Handbook of Contemporary Research*, ed. A. von Stechow and D. Wunderlich. Berlin: Walter de Gruyter pp. 539–50.

Kratzer, Angelika. 2012. *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press.

Lassiter, D. 2015. "Epistemic Comparison, Models of Uncertainty, and the Disjunction Puzzle." *Journal of Semantics* 32(4):649–684.

Lassiter, Daniel. 2010. Gradable Epistemic Modals, Probability, and Scale Structure. In *Proceedings of SALT 20*.

Lassiter, Daniel. 2011. Measurement and Modality: The Scalar Basis of Modal Semantics PhD thesis New York University.

Lauwers, Luc. 2009. "The Uniform Distributions Puzzle." Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1393712.

Leite, Adam. 2010. Fallibilism. In *A Companion to Epistemology, Second Edition*, ed. Jonathan Dancy, Ernest Sosa and Matthias Steup. Blackwell.

Lewis, David. 1980. Index, Context, and Content. In *Philosophy and Grammar*, ed. Stig Kanger and Sven Öhman. Reidel pp. 79–100.

Lewis, David. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74(4):549–567.

Lewis, David K. 1973. *Counterfactuals*. Blackwell.

Littlejohn, Clayton. 2011. "Concessive Knowledge Attributions and Fallibilism." *Philosophy and Phenomenological Research* 83(3):603–619.

Lord, Errol. 2010. "Having Reasons and the Factoring Account." *Philosophical Studies* 149(3):283–296.

Lord, Errol. 2015. Acting for the Right Reasons, Abilities, and Obligation. In *Oxford Studies in Metaethics, vol. 10*, ed. Russ Shafer-Landau. Oxford University Press.

MacFarlane, John. 2011. Epistemic Modals are Assessment Sensitive. In *Epistemic Modality*, ed. Andy Egan and Brian Weatherson. Oxford University Press.

MacFarlane, John. 2014. *Assessment Sensitivity*. Oxford: Oxford University Press.

Marushak, Adam and James R. Shaw. ms. "Epistemics and Emotives." Unpublished manuscript.

McDowell, John. 1994. *Mind and World*. Cambridge: Harvard University Press.

Moore, G. E. 1962. *Commonplace Book, 1919-1953*. New York, Macmillan.

Moss, Sarah. 2013. "Epistemology Formalized." *Philosophical Review* 122(1):1–43.

Moss, Sarah. 2015. "On the Semantics and Pragmatics of Epistemic Vocabulary." *Semantics and Pragmatics* 8(5):1–81.

Neta, Ram. 2008*a*. "What Evidence Do You Have?" *British Journal for the Philosophy of Science* 59(1):89–119.

Neta, Ram. 2008*b*. "What Evidence Do You Have?" *British Journal for the Philosophy of Science* 59(1):89–119.

Neta, Ram. 2011. "A Refutation of Cartesian Fallibilism." *Noûs* 45(4):658–695.

Ninan, Dilip. 2010. "Semantics and the Objects of Assertion." *Linguistics and Philosophy* 33(5):355–380.

Ninan, Dilip. forthcoming. "Relational Semantics and Domain Semantics for Epistemic Modals." *Journal of Philosophical Logic* pp. 1–16.

Papafragou, Anna. 2002. Modality and Theory of Mind: Perspectives from Language Development and Autism. In *Modality and its Interaction with the Verbal System*, ed. S. Barbiers, F. Beukema and W. van der Wurff. Benjamins.

Parfit, Derek. 2011. *On What Matters: Volume One*. Oxford University Press.

Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.

Price, Huw. 1983. "Does 'Probably' Modify Sense?" *Australasian Journal of Philosophy* 61(4):396–408.

Pruss, Alexander R. 2014. "Infinitesimals Are Too Small for Countably Infinite Fair Lotteries." *Synthese* 191(6):1051–1057.

Pryor, James. 2000. "The Skeptic and the Dogmatist." *Noûs* 34(4):517–549.

Quine, W. V. 1960. *Word and Object*. MIT Press.

Rabasca, Michael. 2012. "*Payton v. New York*: Is 'Reason to Believe' Probable Cause or a Lesser Standard?" *Seton Hall Circuit Review* 5(2):437–459.

Reed, Baron. 2002. "How to Think About Fallibilism." *Philosophical Studies* 107(2):143–157.

Reed, Baron. 2012. "Fallibilism." *Philosophy Compass* 7(9):585–596.

Reed, Baron. 2013. "Fallibilism, Epistemic Possibility, and Epistemic Agency." *Philosophical Issues* 23(1):40–69.

Reisner, Andrew and Asbjorn Steglich-Petersen, eds. 2011. *Reasons for Belief*. Cambridge University Press.

Rothschild, Daniel. 2012. "Expressing Credences." *Proceedings of the Aristotelian Society* 112(1pt.1):99–114.

Rysiew, Patrick. 2001. "The Context-Sensitivity of Knowledge Attributions." *Noûs* 35(4):477–514.

Scanlon, T.M. 1998. *What We Owe to Each Other*. Belknap Press of Harvard University Press.

Scanlon, T.M. 2014. *Being Realistic About Reasons*. Oxford University Press.

Schroeder, Mark. 2008. "Having Reasons." *Philosophical Studies* 139(1):57–71.

Schroeder, Mark. 2011. What Does It Take to 'Have' a Reason? In *Reasons for Belief*, ed. Andrew Reisner and Asbjorn Steglich-Petersen. Cambridge University Press pp. 201–22.

Silk, Alex. 2017. "How to Embed an Epistemic Modal: Attitude Problems and Other Defects of Character." *Philosophical Studies* 174(7):1773–1799.

Skorupski, John. 2010. *The Domain of Reasons*. Oxford University Press.

Sobel, David and Steven Wall, eds. 2009. *Reasons for Action*. Cambridge University Press.

Sorensen, Roy. 2009. "Meta-agnosticism: Higher Order Epistemic Possibility." *Mind* 118:777–84.

Stanley, Jason. 2005. "Fallibilism and Concessive Knowledge Attributions." *Analysis* 65(2):126–31.

Stanley, Jason and Zoltán Gendler Szabó. 2000. "On Quantifier Domain Restriction." *Mind and Language* 15(2&3):219–61.

Stojnić, Una. 2016. "One's *Modus Ponens*: Modality, Coherence and Logic." *Philosophy and Phenomenological Research* 93(3).

Stojnić, Una. forthcoming. "Content in a Dynamic Context." *Noûs* .

Suzuki, Satoru. 2013. Epistemic Modals, Qualitative Probability, and Nonstandard Probability. In *Proceedings of the 19th Amsterdam Colloquium*.

Swanson, Eric. 2006. Interactions with Context. PhD thesis Massachusetts Institute of Technology.

Swanson, Eric. 2011. How Not to Theorize about the Language of Subjective Uncertainty. In *Epistemic Modality*, ed. Andy Egan and Brian Weatherson. Oxford University Press.

Swanson, Eric. 2016. "The Application of Constraint Semantics to the Language of Subjective Uncertainty." *Journal of Philosophical Logic* 45(2):121–146.

von Fintel, Kai. 1994. "Restrictions on Quantifier Domains.".

von Fintel, Kai and Anthony Gillies. 2008. "CIA Leaks." *Philosophical Review* 177(1):77–98.

von Fintel, Kai and Anthony Gillies. 2011. "Might" Made Right. In *Epistemic Modality*, ed. Andy Egan and Brian Weatherson. Oxford University Press.

von Fintel, Kai and Irene Heim. ms. "Intensional Semantics." Unpublished manuscript.

Willer, Malte. 2012. "A Remark on Iffy Oughts." *Journal of Philosophy* 109(7):449–461.

Willer, Malte. 2013. "Dynamics of Epistemic Modality." *Philosophical Review* 122(1):45–92.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.

Williamson, Timothy. 2007. "How Probable is an Infinite Sequence of Heads?" *Analysis* 67(295):173–180.

Worsnip, Alex. 2015. "Possibly False Knowledge." *Journal of Philosophy* 112(5):225–246.

Yalcin, Seth. 2007. "Epistemic Modals." *Mind* 116(464):983–1026.

Yalcin, Seth. 2010. "Probability Operators." *Philosophy Compass* 5/11:916–937.

Yalcin, Seth. 2011. Nonfactualism about Epistemic Modality. In *Epistemic Modality*, ed. Andy Egan and Brian Weatherson. Oxford University Press.

Yalcin, Seth. 2012*a*. "Bayesian Expressivism." *Proceedings of the Aristotelian Society* 112(2):123–160.

Yalcin, Seth. 2012*b*. "A Counterexample to Modus Tollens." *Journal of Philosophical Logic* 41(6):1001–1024.

Yalcin, Seth. 2014. Semantics and Metasemantics in the Context of Generative Grammar. In *Metasemantics: New Essays on the Foundations of Meaning*, ed. Alexis Burgess and Brett Sherman. Oxford University Press pp. 17–54.

Yanovich, I. 2014. "Standard Contextualism Strikes Back." *Journal of Semantics* 31(1):67–114.