Proceedings of SALT 24: 597–618, 2014

The weakness of *must*: In defense of a Mantra*

Daniel Lassiter
Stanford University

Abstract Many linguist have claimed that *must*'s meaning is weaker than epistemic necessity—a claim dubbed "the Mantra" in an influential recent paper by von Fintel & Gillies (2008). von Fintel & Gillies argue that the Mantra is false, and that the intuitions that have driven it can be accounted for by appealing to evidential meaning: *must* requires that the proposition it embeds is true and maximally certain, but also known only by indirect means. I show that von Fintel & Gillies do not provide a compelling argument against the Mantra, and that their theory of evidential meaning, while promising in certain respects, also has serious empirical and conceptual problems. In addition, a variety of corpus examples indicate that speakers who assert *must p* are not always maximally confident in the truth of *p*. As an alternative, I re-implement von Fintel & Gillies' theory of indirect evidentiality in a probabilistic, Mantra-compatible framework. Ultimately, both sides of the debate are partly right: *must* is weak in several respects, but it also encodes an indirect evidential meaning.

Keywords: epistemic modality, doxastic modality, evidentiality, inference, probability

1 Introduction

There is a Mantra, for decades repeated mindlessly by researchers in modal semantics: "*Must* is weak". So argue von Fintel & Gillies (2010) in reference to a line of thought going back to Karttunen (1972), who writes:

> There is a striking difference between the logical necessity operator and words like *must*. Consider the two expressions in (1) and (2):
>
> (1) John must have left. (2) John has left.
>
> In any of the standard modal logics, $\Box p$ is a stronger expression than *p*. However, there is an inverse relation between [these] two sentences ... In stating (1), the speaker indicates that he has no first-hand evidence about John's departure, and neither has it been reported to him by trustworthy sources. Instead, (1) seems to say that

* Thanks to Philippe Schlenker, Brett Sherman, Lelia Glass, Matthew Stone, Anastasia Giannikidou, three *SALT 24* reviewers, and audiences at *SALT 24* and the CSLI workshop *Perspectives on Modality*.

> the truth of *John has left* in some way logically follows from other facts the speaker knows and some reasonable assumptions that he is willing to entertain. A man who has seen John leave or has read about it in the newspaper would not ordinarily assert (1), since he is in the position to make the stronger claim in (2).[1]

As von Fintel & Gillies (2010, henceforth "vFG") note, many other linguists have repeated this claim in various forms. Kratzer (1991) formalizes the weakness intuition by making *must*, in effect, a quantifier over a maximally normal subset of the epistemically possible worlds ($\mathcal{E}$); thus *must p* is compatible with there being $\neg p$-worlds in $\mathcal{E}$ as long as they are not maximally normal. This idea has been used in many domains. For example, Giannakidou (1999) uses this feature to explain why certain NPIs are licensed by Greek counterpart of *must* in its deontic *and* epistemic interpretation—though, crucially, not in its alethic interpretation.

However, as Palmer (1979: 59) points out, the intutions supporting the "weakness" claim are confounded with indirectness of evidence. Indeed Karttunen hints at this—"... **no first-hand evidence** about John's departure ..." (my emphasis) —but Palmer's diagnosis is more explicit: "It is the notion of deduction or inference from known facts that is the essential feature of *must*, not just the strength of commitment by the speaker." vFG agree but go further, arguing that *must p* conveys indirectness of evidence but does not convey weakness of any kind: it is false if any $\neg p$-worlds are epistemically possible. They marshal an impressive variety of arguments for the "strength" claim, adding a very interesting novel formalization of indirect evidential meaning which supports the claimed orthogonality of indirectness and weakness.

To illustrate the empirical import of this debate, consider some examples of *must* from the message boards of the genealogy website Ancestry.com.

(3) I have learned that Flora & Dr Alexander had also a brother Dr Murdock MacLeod & a sister Margaret. Margaret married Kenneth MacLeod of Ebost on the Isle of Skye. Must have been about 1820.

(4) My great grandfather John HARTMAN married a lady by the name of Margaret KESSLER. John died in Colorado 1 Oct 1896. His wife was not in his will so must have died before that time. ... Does anybody know anything about this couple?

In (3), the example seems to convey that the conclusion marked by *must* was made on the basis of fragmentary historical records, and not, for example, an explicit report in an authoritative history book. In (4), the evidence is explicitly specified: the author infers that the wife must have died before the husband did on the ground that she

1 Karttunen 1972: 11-12. I've changed example numbering, and updated notation for the necessity operator from 'L' to the now more common '$\Box$'.

was not in his will. Ancestry.com users frequently provide an explicit specification of the evidence used to arrive at a *must* conclusion (see Stone 1994).

Is *must* "weak" or "strong" in these examples? We can ask this question in several different strength-related questions about this example. First, is it plausible that the author of (4), who explicitly acknowledges having very little evidence relevant to the conclusion drawn, is maximally confident in the truth of *The wife died before that time*? Second, is it plausible that the author intends, by marking this proposition with *must*, to **convey** maximal confidence in its truth? Third, does the **truth** of the author's *must*-statement depend on whether, hidden in the mists of time, the wife's absence in the will had a different explanation (say, a marital split)? Or does it depend only on whether the inference that she had died is a compelling one, given the evidence that is available to the author?

As this discussion suggests, the "strength" vs. "weakness" issue should be separated into several components—by my count, at least three distinct questions.

- **Pragmatic strength**: To what extent does a speaker who asserts *must p* take on a commitment to the truth of *p*? Relatedly, to what extent is someone who asserts *must p* blameworthy if *p* turns out to be false?
- **Semantic strength 1** (doxastic status): What does *must p* entail about the status of *p* in $\mathcal{E}$, and/or about its probability? For example, does *must p* entail *It is absolutely certain that p* (or *It is almost certain, likely*, ...)?
- **Semantic strength 2** (veridicity): Does *must p* entail *p*?

vFG argue that *must* is strong on all three dimensions: *must p* entails *p*; it entails maximal confidence in *p* (i.e., *p is absolutely certain*); and a speaker who uses it is maximally committed to the truth of *p*, for example, in that she is blameworthy if *p* subsequently turns out to be false. In this paper I argue that the Mantra is basically right: *must* is not maximal in either the doxastic or pragmatic senses. Cooperative speakers frequently use *must* to mark propositions that they are not maximally certain of, and of which they do not intend to take on a maximally strong pragmatic commitment. Examples (3) and (4) already make a strong case for these claims, since the authors of these texts do not appear to be maximally committed to the truth of the conclusions that they are describing. On the veridicity question, I do not take a strong stand in this paper due to the difficulty of getting clear truth-value judgments about *must* statements (von Fintel & Gillies 2008; Yalcin 2011; MacFarlane 2014). However, I will suggest below that vFG's arguments for veridicity are not compelling.

In section 4 I propose a hybrid theory which treats (in)directness of evidence and degrees of confidence separately. This theory builds conceptually on vFG's account, but combines it with a probabilistic formalism which is better suited for modeling

induction than vFG's deductive approach. This framework allows us to capture the best of Karttunen's and vFG's insights by treating *must* as both indirect and weak.

2 vFG's arguments against weakness

In this and the following section I respond to vFG's many-pronged assault. None of these arguments provide compelling reason to reject the Mantra, though some are convincing as applied to weak theories that do not also encode indirectness.

2.1 Argument 1: *must* is not always weak

vFG emphasize the distinction between indirectness and weakness: conclusions derived from indirect evidence can be maximally strong, and *must* can be used to signal indirectness without even hinting at uncertainty—for instance, in proofs.

(5) x is prime. x is even. x must equal 2.

To my knowledge, though, no "Mantrista" has claimed that *must* **entails** a lack of certainty. For example, Kratzer (1991) treats *must p* as true whenever p is true in all maximally normal worlds in $\mathcal{E}$—but is silent on p's status in non-maximal worlds.

We can, however, extract a related argument against a Kratzer-style weak theory. Suppose that *must p* entails only that p is true in the maximally normal worlds in $\mathcal{E}$. Then, if English has some expression which denotes universal quantification over $\mathcal{E}$, *must p* should be associated with a quantity implicature to the effect that p is not true in **all** of the worlds in $\mathcal{E}$. Suppose that *certainly* is such an expression. Then (6a) should arise sometimes (though by no means always), just like (6b).

(6) a. It must be raining. $\rightsquigarrow$ It's not certainly raining.

b. All of my graduate students are happy. $\rightsquigarrow$ Not all of my students are happy.

If this is a good theory of *certainly*, the absence of an implicature may be a problem for Kratzer's theory: we do not infer from the speaker's failure to conclude (5) with *x certainly equals 2* that the speaker is unsure about basic arithmetical facts. If this argument is convincing, then, *must* cannot be **merely** a non-maximal epistemic modal.

However, the argument from implicature does not apply to a theory which combines weakness with an evidential component. The problem would return if there were a word *certainly′* expressing indirectness **and** universal quantification over $\mathcal{E}$; but the only plausible candidate is *must* itself on vFG's analysis. Since the existence of such an item is exactly what is in dispute, the occasional lack of an inference to lowered confidence tells us nothing.

On the hybrid analysis described below, English allows us to express maximal confidence (using *certainly*) or high-but-not-necessarily-maximal confidence with an evidential signal (using *must*). A speaker's choice to employ *must* could be explained by a desire to emphasize indirectness of evidence, even under full confidence.

2.2 Argument 2: *must* is never weak

vFG give a number of subarguments which target, variously, veridicity, doxastic strength, and degree of pragmatic commitment. None of these arguments gives us a strong reason to reject the Mantra in its most general form.

Subargument 1: Inferencing. Consider the following argument:

(7) "If p, must q. p. Therefore, q."

This feels like a pretty good argument. According to vFG, the intuition that it is provides powerful evidence for one plank of their account, namely, the veridicity of *must*. The reason is that, if *must p* did not entail p, the argument would be logically invalid. For example, on Kratzer's (1991) semantics, *If p, must q* is true as long as q is true in all of the maximally normal worlds in $\mathcal{E} \cap p$. But the actual world might be non-maximal, and so the fact that p is true there is consistent with $\neg q$ being true in the actual world as well.

The argument goes through, but it relies on a questionable background theory of intuitive reasoning: that argument strength intuitions are directly related to deductive validity, and that an argument that is not deductively valid should be perceived as weak. This is a traditional position, and has been argued for prominently in cognitive psychology by Rips (1994). On a theory of this type, we expect gut feelings of argument strength to come in three discrete flavors—"valid", "invalid but consistent", and "inconsistent"—and the distribution of these intuitions should track deductive logic except when people are making reasoning errors.

This position has largely been abandoned in psychology, however. Even Rips (2002) has subsequently abandoned the "imperialist" project of using deductive logic to account for all intuitive reasoning, arguing that we need gradient measures of inductive strength, in addition to deductive validity, in order to account for the full range of argument strength judgments. The best available theory at present is that inductive strength is closely related (perhaps identical) to the **conditional probability** of a conclusion given the premises (Osherson, Smith, Wilkie, Lopez & Shafir 1990; Oaksford & Chater 2001, 2007; Over 2009; Lassiter & Goodman 2015). On this analysis, if the conditional probability of the conclusion given the premises is high, the argument is felt as strong; if it is low, it is felt as weak. Deductive logic remains relevant, since deductively valid arguments are maximally strong

(conditional probability 1) and conclusions which contradict their premises are maximally weak (conditional probability 0). However, it is possible for an argument to be very strong without being deductively valid, if the conditional probability of its conclusion given its premises is close to 1.

On the hybrid, probabilistic theory to be proposed below, the conclusion of (7) is guaranteed to have high probability on the assumption that the premises are true. The first premise requires (inter alia) that $P(q|p)$ is high, and the second that p is true. Combining these two sources of information, we find that $P(q)$ is high, and so the argument is felt as strong.[2] In general, once we have given up the incorrect assumption that argument strength intuitions give us a direct line into deductive validity, there are many possible semantic proposals that could guarantee (7) to be intuitively strong, without making *must* veridical or doxastically maximal.

Subargument 2: *Must* and weak necessity operators. A second attempt to show that *must* is (doxastically) strong involves the observation that (8) is contradictory.

(8) #It must be raining, but perhaps it is not raining.

If *perhaps* expresses existential quantification over $\mathcal{E}$, this is unexpected according to the Mantra, since *must p* does not exclude the presence of $\neg p$-worlds in $\mathcal{E}$. However, if *must* expresses universal quantification over $\mathcal{E}$, the infelicity of (8) is expected.

Unfortunately, this argument misses the target completely. Weak theories, as a rule, deny vFG's assumption that *perhaps* (*might*, *maybe*, *possible*, etc.) expresses existential quantification over $\mathcal{E}$. Instead they treat *perhaps* as *must*'s dual: *perhaps* $p \equiv \neg must \neg p$. On this assumption, (8) is ruled out as infelicitous no matter what *must* means, since it expresses the conjunction of *must p* and $\neg$*must p*.

vFG acknowledge this point briefly in a footnote, responding as follows: "There *are* strong necessity epistemic modals. So pick one and take its dual (e.g., *there is a vanishingly small chance that*). It'll be horrible when paired with *must* in examples like [(8)], we promise" (p.365, fn. 25). In fact it is not too difficult to locate naturally-occurring examples which have more or less this form: here are a few.

(9) There's one missing pepper on the ground a few feet away. A closer look reveals it has been chewed by something. I wouldn't put it past that pesky blue jay to have teeth but then I think it is unlikely he does. It must be a squirrel. What else can get onto a second floor balcony?

The author of the blog post quoted in (9) has pretty shaky ornithological credentials, but their status as an expert speaker of English is not in doubt. This author uses *must* to mark the proposition *It is a squirrel that chewed the pepper*, which contradicts a

2 I am skirting over important issues about how precisely the information in the conditional premise is taken up; see, e.g., Oaksford & Chater 2007.

proposition previously indicated to be possible but unlikely (*the blue jay has teeth* and, by implication, *the blue jay chewed the pepper*).

While the proposition that would render (9) a direct counter-example to vFG's analysis is merely implicit, its counterpart in (10) is fully explicit.

(10) I refuse to believe that this one game, Lost Planet 2 DX11, which was previously 100% stable remember, is crashing because my overclock is unstable .... It's not impossible, granted, but IMO it is highly unlikely. There *must* be some other cause.

Here, *It's not impossible [that Lost Planet 2 DX11 is crashing because my overclock is unstable], but it's highly unlikely* is immediately followed by *There must be some other cause*. This is not compatible with vFG's theory, but it is consistent with a doxastically weak semantics.[3]

These examples indicate that vFG's promise—that *must p* will be intuitively infelicitous when conjoined with any expression hinting that $\neg p$ is epistemically possible—was too hasty: such conjunctions are possible and attested. What is more, these examples directly falsify vFG's claim that *must* is a universal quantifier over $\mathcal{E}$, and provide strong support for the Mantra.[4]

Subargument 3: Distancing. Recall the distinction between three kinds of strength from section 1: veridicity, doxastic strength, and pragmatic commitment. The contrast between doxastic strength and veridicity should be apparent: someone could be maximally certain of the truth of a false sentence, or less than fully certain of a the truth of a veridical statement. What is less clear is how pragmatic commitment relates to either variety of semantic strength. This point is important in the current context because many of the arguments for and against the Mantra invoke pragmatic commitment, rather than directly targeting doxastic strength.

One such argument involves whether it is felicitous for someone who utters *p*, *might p*, *must p*, etc. to distance herself from *p* if it turns out to be untrue. vFG argue that such distancing is not acceptable with *must*, as in (11).

(11) a. **Alex**: It must be raining.

b. **Billy**: [opens curtains] No it isn't. You were wrong.

3 Note, however, that these examples place constraints on the meanings of the other epistemic expressions in this. For example, to account for (10) we must assume (quite plausibly) that *p is impossible* requires a higher degree of confidence in $\neg p$ than *must* $\neg p$ does.

4 It has been suggested to me that there is some kind of modal domain shifting going on in the midst of these examples. There is no way to exclude this possibility, but without detailed constraints on the inclusion and exclusion of worlds from $\mathcal{E}$ such a theory borders on vacuity: any two constraints on epistemic states could consistently be placed side-by-side if we allow unconstrained domain shifting. It would, however, be very interesting to see an empirically predictive theory of this type spelled out.

c. **Alex**: # I was not! Look, I didn't say it *was* raining. I only said it *must* be raining. Stop picking on me!

Alex's response could, however, be appropriate with *might*, *ought* or *plausible*.

In vFG's judgment, "saying *I only said it must be raining* is as bizarre as *I only said I ate all of the cookies*. But if, as the Mantra maintains, *must* is not located at the very top of the scale of epistemic strength, one would expect *only* and *must* to combine like old friends" (von Fintel & Gillies 2010: 366). This argument is not about veridicity: if we substitute the non-veridical *I'm completely sure it's raining* in (11), the oddness remains. Rather, the argument involves doxastic strength (via the interaction with *only*) and pragmatic commitment (via intuitions of the appropriateness of distancing).

There are various objections that one could make to the distancing argument: the intuition that (11c) is odd is neither clear nor universal, and the "old friends" argument assumes a simplistic picture of what *only* means (see Beaver & Clark 2008). But the most serious problem is that the oddness of (11c) depends on features of this specific example. In a different context, a response along these lines could be used appropriately to emphasize that her statement was reasonable given the evidence available. Consider a variant of the genealogy examples given in section 1:

(12) a. **Alex**: John died in Colorado 1 Oct 1896. His wife was not in his will so must have died before that time.

b. **Billy** [consults her private archive, which no one else knows about]: No she hadn't—you had not consulted a detailed source on their lives which you didn't know about, and which only I have access to. She left him in 1893, and was living in New Orleans on that date. You were wrong.

c. **Alex**: I was not! Look, I didn't say she *had* died—I had no way to know that for sure. I only said she *must have* died, and that was a totally reasonable conclusion given what was available to me.

Unlike (11c), Alex's response in (12c) seems reasonable. I suspect that the difference in felicity between (11) and (12) involves whether it would have been easy for the speaker to access evidence that would be decisive about the truth of the embedded sentence (see Hacking 1967; DeRose 1991). In (11), where Alex could easily have resolved the question by opening the curtains, it seems appropriate for Billy to blame her for making a mistake, and odd for her to refuse to accept blame. But when it would have been difficult or impossible for the Alex to access decisive evidence, *I only said must p* can be used to convey that Billy's attempt to assign blame is unreasonable: "Based on what was available at the time, I asserted—quite reasonably—that the evidence strongly favored the conclusion that *p*".

2.3 Argument 3: A strong semantics makes available an attractive account of evidential meaning.

As noted above, several earlier authors indicate a connection between *must* and reasoning from indirect evidence (Karttunen 1972; Palmer 1979). However, vFG seem to be the first to suggest that the **only** relevant factor is (in)directness, and that strength of commitment is irrelevant to the meaning of *must*. To reinforce this claim, vFG propose a positive account on which *must* is a universal quantifier over $\mathcal{E}$, which also encodes a presupposition of indirect evidence. Part of the attraction of the theory is that it encodes *must*'s requirement of indirect evidence in a way that is consistent with its veridicity and maximal doxastic and pragmatic strength. However, this positive theory encounters both empirical and conceptual problems. Fortunately, these problems also point the way for an improved semantics—though one that is compatible with the dreaded Mantra.

In vFG's account, there is a set of propositions $\mathcal{K}$ (the **kernel**) each of which are known through direct experience to be true. They assume that the set of epistemically possible worlds $\mathcal{E}$ is the deductive closure of $\mathcal{K}$, i.e., $\mathcal{E} = \bigcap \mathcal{K}$; this assumption makes sense if we allow that whatever is known is known either through some kind of direct perception or through inference from direct perception. The propositions known indirectly, $\mathcal{I}$, are those that are entailed by $\mathcal{E}$ but not in $\mathcal{K}$:

$$\mathcal{I} = \{p | \mathcal{E} \subseteq p\} - \mathcal{K}.$$

Crucially, entailment is the only kind of inferencing that this account permits. vFG then define truth-conditions for a simple epistemic *must* sentence:

(13) *must p* $\begin{cases} \text{is a presupposition failure if } p \in \mathcal{K} \text{ (if } p \text{ is known directly);} \\ \text{otherwise, is true iff } \mathcal{E} \subseteq p \text{ (i.e., if } \mathcal{E} \text{ entails } p\text{).} \end{cases}$

This definition captures indirectness via the presupposition that $p \notin \mathcal{K}$; when this presupposition is satisfied, it asserts that p is known. As a secondary inference, if *must p* is not a presupposition failure, we can infer that $p \in \mathcal{I}$, the set of propositions that are known indirectly. Extrapolating somewhat from vFG's explicit discussion, it appears that this semantics derives the three planks of their account as follows.

Doxastic maximality: First, we must assume that, when something is known, it is believed with maximal confidence. Second, we must assume that deduction preserves maximal confidence. On these assumptions, something that is known via deduction from known premises is known with maximal confidence.

Maximal pragmatic strength: Asserting something that entails maximal belief in p is inappropriate unless one is maximally committed to the truth of p.

Veridicity: "You can't have direct information that *P* unless it is the case that *P*. ... So our modal bases will be reflexive" (von Fintel & Gillies 2010: 371).

To illustrate the application of this semantics to a concrete case, vFG introduce Billy, who has just seen people coming into her office with wet raincoats on. She utters "It must be raining." The semantics predicts that this utterance is a presupposition failure if Billy has directly observed that it is raining, and that, otherwise, it is true just in case Billy has access to some kernel $\mathcal{K}$ of true and directly observed propositions such that $\bigcap\mathcal{K}$ entails that it's raining. Specifically, vFG suggest that Billy's utterance is appropriate and true if her kernel contains at least the following:

(14) $\mathcal{K}_{\text{Billy}} \supseteq \{$**people are coming in with wet raincoats**, **people only come in with wet raincoats when it's raining**$\}$

The deductive closure of any such $\mathcal{K}$ must also include the proposition **it's raining**.

The basic problem with this proposal is that it cannot cope with the frequent use of *must* to report inductive inferences. In the genealogical examples in (3) and (4), for instance, *must* is used to mark what look like "inferences to the best explanation": statements which are reasonable to believe because their truth would readily **explain** the available evidence, but which are not entailed by anything known to the author. In these examples it is difficult to even imagine a set of propositions that would entail the truth of the statement and which the author could plausibly have observed directly. It is also implausible to suppose that the authors have an **erroneous** belief that they have access to the required set of propositions: experienced genealogists are well-aware of the epistemic status of their inferences. (More in section 3.)

But vFG's problems with induction are even closer to home: it should be inappropriate for Billy to use *must* even in the scenario that vFG focus on. The reason is that Billy's kernel crucially contains the "directly known" proposition **people only come in with wet raincoats when it's raining**. This is, of course, logically equivalent to the negative existential **People never come in with wet raincoats when it's not raining**. How could Billy acquire **direct** knowledge of the truth of this negative existential proposition? The only way to do so, it seems, would be to actually witness all relevant situations, and check that none of them are situations in which (a) it is not raining, and (b) someone comes in with a wet raincoat. The set of relevant situations is presumably infinite, and so this is not possible even in principle: no amount of experience could grant Billy direct knowledge of the non-existence of a situation type. But then Billy's utterance of *It must be raining* is false: *It is raining* is not entailed by what she knows from direct experience.

This is not a minor issue that can be attributed to a poorly-chosen example. Instead, it reveals a deep conceptual problem—and one which seems inevitable given (a) vFG's assumption that "inference = deduction", and (b) the fact that *must* is frequently used to report inductive and other non-trivial inferences. The kernel-based analysis of the evidential component of *must* is too restrictive to deal with this example, or—as we see in section 3—with other, more naturalistic uses of *must*.

3 Analysis of naturalistic examples

This section discusses a number of additional corpus-derived examples of epistemic *must*. These illustrate additional empirical problems for vFG's account, but also, on the positive side, point toward features that a successful theory of *must* must have.

Lack of knowledge or direct evidence. Like (3) and (4), the examples to follow were collected from genealogical discussions which are largely devoted to speculation about the lives of unknown, long-dead persons. These discussions attempt to piece together details and dates of births, deaths, marriages, and migrations, on the basis of fragmentary historical records, general reasoning, and hearsay. The authors frequently use *must* to mark inferences which have been made in this fashion. The equation *must* = *known by deduction from direct observation* does not do them justice: genealogists and other historians frequently use *must p* despite clearly being aware that they do not know anything that entails *p*. Consider (15), taken from *The Plymouth Colony Archive Project*'s discussion of the fates of *Mayflower* passengers.

(15) Goodman, John. "Died soon after arrival in the general sickness," ... Goodman was still alive in mid-January 1621 ... , although not in good physical shape. He is listed as one of those who received land in 1623 ... He is not listed among those who were part of the cattle division of 1627, so he must have died by then.

For vFG to account for this example, $\mathcal{K}$ would have to include something like **The only way that a farmer who arrived on the *Mayflower* in 1620, is reported to have died "soon" afterwards, was sick in 1621, and was reported to be alive in 1623, could fail to be in a list of farmers in 1627 is that he died in the meantime**. No one could seriously self-ascribe knowledge of this proposition, direct or otherwise. Alternatively, we could suppose that the authors are uncooperative or flouting conversational norms for pragmatic effect; but neither explanation is plausible here.

A better diagnosis is that the semantic account under consideration is incorrect. These authors are not trying to express maximal confidence in the truth of *Goodman had died by 1627*, nor to take on a maximal commitment to this proposition. Instead, (15) presents Goodman's death as the **best explanation** of the evidence presented. This is indeed a good explanation, and it is rendered highly probable by what is reported; but it is in no way entailed by this evidence.[5]

5 It gets worse. It is well-known among genealogists that the sources on which they rely are often unreliable (Mills 1999). If vFG's semantics were correct, we would expect that competent genealogists would almost never use *must*; and that we, knowing about the possible unreliability of sources, would perceive most instances of *must* in this context as infelicitous. But epistemic *must* is quite frequent in genealogical discussions, and seems to be a natural way to report best-guesses about past events.

More examples like this could easily be adduced. Indeed, similar lessons emerge from the two genealogical examples that we saw in section 1. For instance, example (4)—*His wife was not in his will so must have died before that time*—invokes a line of reasoning which is plausible, but which could hardly be said to be foolproof. Again, we would have to be quite uncharitable to the author in order to explain the use of *must* here if vFG's account were correct—supposing that the author does not realize that marriages can end for reasons other than death.

Explicit disavowal of direct evidence. A clear reason to doubt that entailment from direct evidence is a requirement of *must* is that speakers sometimes use *must p* while specifically denying possession of direct evidence for *p*.

(16) Almost certainly the site must have been inhabited well before that time, but in a place where virtually every square inch of land has been built and rebuilt upon many times over the centuries, positive evidence is most difficult to uncover ...

(17) Probably this must have been done before, but I couldn't find enough information on this in the ISIS doc & ISIS/GIS community forums.

Co-occurrence with weaker epistemic modals and attitudes. (16) and (17) are additionally interesting because they show *must* co-occuring with the (weaker?) epistemic modals *almost certainly* and *probably*. I have found a number of such examples, including cases involving the clearly weaker modals *perhaps* and *maybe* and the attitude verb *figure*.

(18) [I]n fact, the words we hear as 'pity' can also be translated to mean that when Jesus looked at the man, he 'snorted like a war horse.' Now that's some kind of anger. It's deeply rooted, instinctive even. As perhaps it must have been.

(19) Last August, when they called me and asked whether I'd speak at The Global Leadership Summit held by the Willow Creek Association, I thought maybe there must have been a mix up.

(20) If the handgun was engraved or had some sort of fancier finish then I figured he must be a "pistolero." I might have been wrong but those were my initial impressions.

The authors of (16)–(19) are clearly not, for example, attempting to express the claim that the speaker {perhaps/maybe/ probably/almost certainly} has direct information which entails the proposition embedded by *must*. After all, in (16) and (17) the authors explicitly disavow possession of such evidence. The author of (20) even follows *I figured must p* by acknowledging the possibility of error and his fragile

epistemic basis for concluding *p*: "*I might have been wrong*, but those were *my initial impressions*." These examples cannot be treated as standard modal concord, which is supposed to require strength matching, or as examples of vacuous quantification by the modal with wider scope (Hacquard 2006; Yalcin 2007), since both modals contribute elements of the interpretation.

In these examples the message is more along the lines of "I infer, with a degree of confidence associated with {perhaps/maybe/probably/almost certainly/figure}, that *p*." *Must* apparently retains its indirectness requirement, but cedes control over doxastic strength to the other epistemic expression. While length restrictions prevent me from attempting a detailed treatment of these interactions here, I believe that the threshold-based semantics that I propose below can provide an account of such interactions on the analogy of degree modification in expressions like *somewhat tall*—more or less along the lines of the "modal modification" accounts of such interactions explored by Grosz (2010); Anand & Brasoveanu (2010); Giannakidou & Mari (To appear).

The relevance of statistical reasoning. Example (21)—pieced together from messages and replies on the Ancestry.com message boards—illustrates the possible relevance of statistical reasoning to a speaker's choice to employ *must*. The key point is Author 1's account in (21c) of his earlier choice to employ *must*: "I was only assuming", he explains, and then describes the reasoning behind his assumption.

(21) a. **Author 1**: [Y]our man Lazarus must have sustained injuries at [Buena Vista] by his death date. ...

b. **Author 2**: I check the killed and wounded list for the Battle of Buena Vista and Lazarus wasn't listed under killed and wounded.

c. **Author 1**: Curious. I was only assuming that since Lazarus is listed as dying on March 2, '47, that it was from wounds suffered the week prior (Buena Vista was fought Feb. 22-23). His unit (as with all volunteers at the battle) was certainly in the thick of it. Nevertheless, as we all know, disease took a heavier toll on the troops than actual enemy fire. In my research, though, when I see a death date that close to the battle date, I tend to think that wounds played apart.

Author 1's choice to use *must* in (21a) is not well accounted for by supposing that he thought he knew (directly or otherwise) that Lazarus had been wounded. After all, he explains in (21c), "I was only assuming ..." and "I tend to think ... ". In addition, Author 1 is an expert on this domain—he mentions elsewhere that "I have been researching the battle of Buena Vista for several years"—and he notes that "we all know" that disease was the primary cause of death. It seems unlikely that this author forgot that disease was a possible cause of death when formulating (21a).

The explanation continues with a description of implicit statistical reasoning:

- Most soldiers who died in the war died of disease: $P(\textbf{disease}|\textbf{died})$ is high.
- Also, $P(\textbf{wounded}|\textbf{died})$ is low: most who died had not been wounded.
- However, $P(\textbf{wounded}|\textbf{battle}\&\textbf{died})$ is high: Death *within a week of battle* is usually attributable to battle wounds.
- Since Lazarus died within a week of battle, Lazarus must have been wounded.

Lacking more specific evidence regarding **wounded** when (21a) was formulated, Author 1 chose *must* to mark **wounded** as the **best explanation** of the available data.

4 Abductive and threshold semantics

I now propose a way of formalizing the kind of inference from indirect evidence that vFG correctly identify as the core of *must*'s meaning, but doing it in a way that is compatible with the Mantra and with the naturalistic examples of *must* that were discussed above. Inference—including induction—is encoded in this model as probabilistic reasoning. This model preserves much of vFG's attractive account of evidential meaning, but renders it compatible with the Mantra.

Note: I am **not** claiming that the account to follow is the only possible way to explain the meaning and use of *must*. There are surely other possible theories of *must* that encode an indirectness requirement and make room for induction. For example, some variety of default logic (Reiter 1980) might be made to work. I focus on a probabilistic account because I believe that it coheres best with cutting-edge work in the semantics of epistemic modals and in the cognitive science of inference.[6]

4.1 Question-based probabilistic information dynamics

Probabilistic graphical models (PGMs) and their variants represent the state of the art in representing and reasoning about uncertainty in AI and psychology.[7] In this approach, like other subjective-probability formalisms, information states are

6 On probability in the semantics of epistemic modals, see Yalcin 2005, 2007, 2010; Swanson 2006; Lassiter 2010, 2011, 2014a, To appear, 2014b; Klecha 2012; Moss To appear. For a small sampling of relevant research on inference in cognitive science, see footnote 7 and also Osherson et al. 1990; Glymour 2001; Tenenbaum & Griffiths 2001; Tenenbaum, Griffiths & Kemp 2006; Oaksford & Chater 2007; Kemp & Tenenbaum 2009; Lassiter & Goodman 2015.

7 Artificial Intelligence: Pearl 1988, 2000; Spirtes, Glymour & Scheines 1993; Koller & Friedman 2009; Russell & Norvig 2010, and many more. Psychology: Glymour 2001; Sloman 2005; Gopnik & Schultz 2007; Danks 2014, and—with additions—Tenenbaum, Kemp, Griffiths & Goodman 2011.

represented using a set of epistemically possible worlds $\mathcal{E}$ together with a measure function P which (a) takes propositions to $[0,1]$, (b) assigns 1 to any tautology, and (c) is additive—$P(A \cup B) = P(A) + P(B)$ if $A \cap B = \varnothing$. Learning and inference are modeled as conditionalization. If the probability distribution is $P(\cdot)$ and then B is learned ($P(B) \neq 0$), the distribution is updated to incorporate the information that B is true. Afterwards, for any proposition C, $P(C|B) = P(B \wedge C)/P(B)$.

PGMs add structure to this basic picture in two ways. First, they contain a set of "variables", each of which partitions $\mathcal{E}$. Formally, a variable is simply a Groenendijk & Stokhof 1984 question denotation restricted to $\mathcal{E}$ (see van Rooij 2003).

(22) **Definition** (Variable/Question). A set of propositions ("answers") s.t.

a. All answers in Q are subsets of $\mathcal{E}$.

b. Any two distinct answers to Q are mutually exclusive.

c. Q covers $\mathcal{E}$: one of the answers to Q must be true.

Second, PGMs add "arrows" representing dependencies between variables/questions. If there is an arrow from Q_1 to Q_2, then Q_1 and Q_2 are dependent in the sense that learning the value of one can influence the distribution on the other. Independent variables have the inverse property: if Q_1 and Q_2 are independent then there is no answer to Q_1 which, if it were learned, would lead the learner to modify the probability that she assigns to any answer to Q_2 (and vice versa). (In)dependence assumptions will not play a major role here, but they are critical for many applications.

Building on the PGMs framework, and with inspiration from recent formal models of awareness (Franke & de Jager 2007; de Jager 2009; Yalcin 2011), I assume that agents represent only some of the possible questions—that is, that they are sensitive only to certain issues at a given time. Let the relevant set of questions be $\mathcal{V}$. As a set of question meanings, $\mathcal{V}$ is also a set of variables in the PGMs sense. I assume further that $\mathcal{V}$ is partitioned into two sets: a set $\mathcal{V}_\mathcal{D}$, the variables whose true values have been observed directly, and $\mathcal{V}_\mathcal{I} = \mathcal{V} - \mathcal{V}_\mathcal{D}$, whose values have not been observed directly and must be inferred. This division is directly inspired by vFG's distinction between a "Kernel" $\mathcal{K}$ and a set of indirectly known propositions, but its formal properties are significantly different.

Let P_m denote the relevant probability measure at time t_m. Suppose that, at t_n, the answer to question $\mathcal{Q}$ is directly observed (and no other new information is acquired). These conditions must hold at all times after t_n:

- If $\mathcal{Q}$ is in $\mathcal{V}_\mathcal{I}$ at t_n, it is in $\mathcal{V}_\mathcal{D}$ at t_k for $k > n$.
- There is some answer $q \in \mathcal{Q}$ such that $P_k(q) = 1$ for $k > n$.
- Suppose, after t_n, no other question's answer is directly observed until t_p. Then, for all t_k such that $t_n < t_k < t_p$, and for all $\mathcal{Q}' \in \mathcal{V}$, $P_k(\mathcal{Q}') = P_n(\mathcal{Q}'|\mathcal{Q})$.

The first condition updates the direct/indirect division: when an agent observes the answer to a question about which she has previously made only probabilistic inferences, that question is registered thereafter as one whose answer has been directly observed. The second requires that direct observation is uncertainty-free and that observed answers are not forgotten. (These are helpful but not insignificant simplifications). For instance, if we are representing the question *Is it raining at t_m?* and then observe that it's raining at t_m, then $P(\textbf{rain-at-}t_m) = 1$ for all times thereafter. The third condition requires that information dynamics proceed exclusively by Bayesian update on direct observations, ruling out non-Bayesian update and clairvoyance.

A further notable feature of the third condition is that, at any time t_m, the distribution P_m is equal to the initial distribution P updated by the values of all directly observed variables: $P_m(\mathcal{Q}) = P(\mathcal{Q}|\mathcal{V}_\mathcal{D})$, where $\mathcal{V}_\mathcal{D}$ is the set of all questions whose values have been observed up to time t_m. Given this, we can simply write $P(\cdot|\mathcal{V}_\mathcal{D})$ for the relevant probability distribution at any time, letting the identity of $\mathcal{V}_\mathcal{D}$ take care of the temporal bookkeeping.

4.2 Capturing *must*

Above we discussed a number of examples which presented difficulty for two planks of vFG's proposal: the claim that *must p* entails maximal speaker belief in the truth of p, and the claim that it typically induces a maximal pragmatic commitment on the part of speaker to the truth of p. I argued, essentially following Stone 1994, that the effect of *must* in a variety of naturalistic examples would be better paraphrased "p is the best explanation of the available evidence".

With the PGM-based model just sketched in hand, it is straightforward to write down a semantics for *must* which is closely related to vFG's, but replaces their "entailed by direct information" condition with a "best explanation" condition. The definition assumes that assertions are required to address a Question Under Discussion (QUD) which is determined by the discourse context (Ginzburg 1995a,b; van Kuppevelt 1995; Roberts 2012; Beaver & Clark 2008). An assertion of *must* q presupposes that q addresses the QUD $\mathcal{Q}$—simplifying, that it is an element of (direct answer to) $\mathcal{Q}$. This is added to *must*'s presupposition that $\mathcal{Q}$ is not in $\mathcal{V}_\mathcal{D}$.

(23) *must* q $\begin{cases} \text{is a presupposition failure if } q \notin \mathcal{Q} \text{ or } \mathcal{Q} \in \mathcal{V}_\mathcal{D}; \\ \text{otherwise, is true iff } \forall q' \neq q \in \mathcal{Q} : P(q|\mathcal{V}_\mathcal{D}) > P(q'|\mathcal{V}_\mathcal{D}). \end{cases}$

If we want to make *must* veridical we could add the truth of q as an additional entailment. I do not know if this is the right choice empirically, but it seems to be independent of the other details discussed here.

Bracketing the issue of veridicity, (23) effectively presents q as the answer to question $\mathcal{Q}$ which has the highest probability because it best explains the observed

evidence $\mathcal{V}_\mathcal{D}$ given the background assumptions encoded in P. As an example, consider (4) again. Here the QUD might be ?q =*Did Hartmann's wife die before 1 Oct 1896?*, a partition with two cells: $\{q, \neg q\}$. Then the statement that the speaker makes—*must* q—is predicted to be **felicitous** only if the speaker has not directly observed either q or $\neg q$. Infelicity might result, for example, if the speaker had been present at Hartmann's wife's death, or had received a reliable report from someone who was, or had seen a death certificate. If felicitous, *must* q is true iff q is more probable than $\neg q$, conditional on $\mathcal{V}_\mathcal{D}$.

This is a decent rendition of what (4) conveys, though it seems too weak: we should probably also require that p be *much* more likely than $\neg p$. But the problems go deeper. If the QUD has many possible answers, the most likely could still be very improbable. For example, the question *On what day did Hartmann's wife die?* has many possible answers, all of which might have very low probability. If q = *She died on day 1 Jan 1896* is the most likely, we do not necessarily want *must* q to come out true for this reason: with thousands of possible answers, this condition could hold even if has probability less than 1/1000. However, it seems clear that *must* q should be false when $P(q|\mathcal{V}_\mathcal{D})$ is this low.

We can avoid this issue by strengthening the definition as follows: *must* q requires that q is highly probable, which—as long as the required probability is at least .5—entails that it is the most probable element of any partition that contains it.

(24) *must* q $\begin{cases} \text{is a presupposition failure if } q \notin \mathcal{Q} \text{ or } \mathcal{Q} \in \mathcal{V}_\mathcal{D}; \\ \text{otherwise, is true iff } P(q|\mathcal{V}_\mathcal{D}) > \theta_{must}. \end{cases}$

With $\theta_{must} \geq .5$, the definition in (24) is strictly stronger than the one in (23): it requires that q be the best answer to the QUD, and—when the QUD is highly ramified—the best by far. The definition still leaves much to be said, though: we have not yet said anything about how uncertainty about θ_{must} is resolved, or why it should be $\geq .5$.

Must is in pragmatic competition with a number of other items to describe degrees of uncertainty or (in)directness of evidence. Many of these items are also vague scalar expressions with context-sensitive meanings, such as *(im)plausible*, *(un)likely*, *(un)certain*, *(im)possible*, and—I would argue—*might*. It is reasonable to suppose that whatever pragmatic mechanisms go into estimating the vague and context-sensitive threshold values for scalar adjectives—*tall, short, heavy, warm, hot, full, wet*, etc.–are also recruited for the purpose of estimating the threshold value that it implicated in *must*'s meaning. The closest analogue, on this way of thinking, would be the choice between *warm* and *hot* to describe temperature. Even though their interpretations vary greatly as a function of context and reference class, lexical knowledge plays a partial role in fixing their thresholds: we know that the minimum temperature required for something to be *hot* is greater than that required

to be warm, rather than the reverse. In other words, even though the **values** of θ_{warm} and θ_{hot} are not fixed by our linguistic knowledge, the **ordering** $\theta_{warm} < \theta_{hot}$ is.

While epistemic expressions may have vague and context-sensitive meanings, there are facts about their orderings that are similarly fixed lexically. For example, the scalar epistemic adjectives *probable* and *likely* have context-sensitive meanings that change with the distribution of probabilities among salient alternative sets (Teigen 1988; Windschitl & Wells 1998; Yalcin 2010; Lassiter 2011). In addition, a recent experimental examination of epistemic readings of six adjectives found evidence for the ordering *possible* < *plausible* < *probable* < *likely* < *certain*/*necessary* (Lassiter & Goodman 2015) . The obvious way to enforce the ordering *probable* < *likely* is to make $\theta_{probable} < \theta_{likely}$ a lexical feature of these adjectives, analogous to $\theta_{warm} < \theta_{hot}$. This constraint will ensure that the two thresholds move in lockstep, regardless of the mechanisms and information sources that are recruited to resolve the interpretation further. We can constrain the interpretation of *must* similarly with a lexical ordering relative to other epistemic expressions: for example, θ_{must} is probably required to be greater than θ_{likely}. Perhaps it is also constrained to be less than $\theta_{certain}$.

While there is still much to be said about the way that speakers and listeners narrow in on a more precise interpretation of *must*, the comparison with scalar adjectives raises the hope that a theory of the threshold inference process for scalar adjectives could simply be plugged in to derive reasonable results for *must*. See Lassiter & Goodman 2013, To appear for a recent pragmatic theory of threshold inference which is compatible with the present account.

5 Conclusion

The semantics of *must* is a rich topic with many connections to other epistemic items and to philosophical and psychological theories of knowledge and inference. In this brief paper I have only hinted at possible answers to many of the questions raised by this item's meaning and use. Some important topics for further research include

- a compositional semantics for *must* under attitude verbs like *figure* (20) and in its co-occurrence with weaker epistemic items like *almost certainly, possibly, perhaps* (16-19);
- the connection between *must* and knowledge: is there one? if so, does it mean that knowledge is also "weak" in some sense?;
- empirical and computational investigation of the semantics and pragmatics of vagueness in the meaning of *must*.

But we can't do all of this here, so let's close with the Mantra instead: "*Must* is weak! *Must* is weak! *Must* is weak! *Must* is weak! *Must* is weak! *Must* is weak! ..."

References

Anand, Pranav & Adrian Brasoveanu. 2010. Modal concord as modal modification. In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds.), *Sinn und Bedeutung 14*, 19–36. Universität Wien Institut für Sprachwissenschaft. http://www.univie.ac.at/sub14/proc/grosz.pdf.

Beaver, David & Brady Clark. 2008. *Sense and Sensitivity: How Focus Determines Meaning*. Wiley-Blackwell.

Danks, David. 2014. *Unifying the Mind: Cognitive Representations as Graphical Models*. MIT Press.

DeRose, Keith. 1991. Epistemic possibilities. *The Philosophical Review* 581–605.

von Fintel, Kai & Anthony Gillies. 2008. CIA leaks. *The Philosophical Review* 117(1). 77.

von Fintel, Kai & Anthony Gillies. 2010. Must... stay... strong! *Natural Language Semantics* 18(4). 351–383.

Franke, Michael & Tikitu de Jager. 2007. The relevance of awareness. In Paul Dekker Maria Aloni & Floris Roelofsen (eds.), *Sixteenth Amsterdam Colloquium*, 97–102. ILLC/Department of Philosophy, University of Amsterdam. http://www.illc.uva.nl/AC/AC2007/uploaded_files/proceedings-AC07.pdf.

Giannakidou, Anastasia. 1999. Affective dependencies. *Linguistics and Philosophy* 22(4). 367–421.

Giannakidou, Anastasia & Alda Mari. To appear. Epistemic weakening with future and *must*: Non-veridicality, evidentiality, and partial knowledge. In Joanna Blaszack, Dorota Klimek-Jankowska, Kryzstof Mygdalski & Anastasia Giannikidou (eds.), *Tense, Mood, and Modality: New Perspectives on Old Questions*, University of Chicago Press.

Ginzburg, Jonathan. 1995a. Resolving questions, I. *Linguistics and Philosophy* 18(5). 459–527.

Ginzburg, Jonathan. 1995b. Resolving questions, II. *Linguistics and Philosophy* 18(6). 567–609.

Glymour, Clark N. 2001. *The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology*. MIT press.

Gopnik, Alison & Laura Schultz (eds.). 2007. *Causal Learning: Psychology, Philosophy, and Computation*. Oxford University Press.

Groenendijk, Jeroen & Martin Stokhof. 1984. Studies in the Semantics of Questions and the Pragmatics of Answers. Ph.D. thesis, University of Amsterdam.

Grosz, Patrick. 2010. Grading modality: A new approach to modal concord and its relatives. In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds.), *Sinn und Bedeutung 14*, 185–201. Universität Wien Institut für Sprachwissenschaft. http://www.univie.ac.at/sub14/proc/grosz.pdf.

Hacking, Ian. 1967. Possibility. *The Philosophical Review* 143–168.
Hacquard, Valentine. 2006. Aspects of Modality. Ph.D. Thesis, MIT.
de Jager, Tikitu. 2009. "Now That You Mention It, I Wonder...": Awareness, Attention, Assumption. Ph.D. thesis, University of Amsterdam.
Karttunen, Lauri. 1972. Possible and must. In John Kimball (ed.), *Syntax and semantics*, vol. 1, Seminar Press.
Kemp, Charles & Joshua B. Tenenbaum. 2009. Structured statistical models of inductive reasoning. *Psychological Review* 116(1). 20–58.
Klecha, Peter. 2012. Positive and conditional semantics for gradable modals. In Anna Chernilovskaya Ana Aguilar Guevara & Rick Nouwen (eds.), *Sinn und Bedeutung 16*, vol. 2, 363–376. MIT Press. http://mitwpl.mit.edu/open/sub16/.
Koller, Daphne & Nir Friedman. 2009. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press.
Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An International Handbook of Contemporary Research*, 639–650. de Gruyter.
van Kuppevelt, Jan. 1995. Discourse structure, topicality and questioning. *Journal of Linguistics* 31(01). 109–147.
Lassiter, Daniel. 2010. Gradable epistemic modals, probability, and scale structure. In Nan Li & David Lutz (eds.), *Semantics & Linguistic Theory (SALT) 20*, 197–215. CLC Publications.
Lassiter, Daniel. 2011. Measurement and Modality: The Scalar Basis of Modal Semantics. Ph.D. thesis, New York University.
Lassiter, Daniel. 2014a. Epistemic comparison, models of uncertainty, and the disjunction puzzle. *Journal of Semantics* Advance Access 1–36. doi:10.1093/jos/ffu008.
Lassiter, Daniel. 2014b. Modality, scale structure, and scalar reasoning. *Pacific Philosophical Quarterly* 95(4). 461–490.
Lassiter, Daniel. To appear. *Measurement and Modality: The Scalar Basis of Modal Semantics*. Oxford University Press.
Lassiter, Daniel & Noah D. Goodman. 2013. Context, scale structure, and statistics in the interpretation of positive-form adjectives. In Todd Snider (ed.), *Semantics & Linguistic Theory (SALT) 23*, 587–610. CLC Publications.
Lassiter, Daniel & Noah D. Goodman. 2015. How many kinds of reasoning? Inference, probability, and natural language semantics. *Cognition* 136. 123–134.
Lassiter, Daniel & Noah D. Goodman. To appear. Adjectival vagueness in a Bayesian model of interpretation. *Synthese*.
MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and its Applications*. Oxford University Press.
Mills, Elizabeth S. 1999. Working with historical evidence: Genealogical principles

and standards. *National Genealogical Society Quarterly* 87(3). 165–84.
Moss, Sarah. To appear. The semantics and pragmatics of epistemic modals. *Semantics and Pragmatics*.
Oaksford, Mike & Nick Chater. 2001. The probabilistic approach to human reasoning. *Trends in Cognitive Sciences* 5(8). 349–357.
Oaksford, Mike & Nick Chater. 2007. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford University Press.
Osherson, Daniel N., Edward E. Smith, Ormond Wilkie, Alejandro Lopez & Eldar Shafir. 1990. Category-based induction. *Psychological Review* 97(2). 185.
Over, David E. 2009. New paradigm psychology of reasoning. *Thinking and Reasoning* 15(4). 431–438. doi:10.1080/13546780903266188.
Palmer, Frank Robert. 1979. *Modality and the English Modals*. Longman London.
Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
Pearl, Judea. 2000. *Causality: Models, Reasoning and Inference*. Cambridge University Press.
Reiter, Raymond. 1980. A logic for default reasoning. *Artificial Intelligence* 13(1). 81–132.
Rips, Lance. 1994. *The Psychology of Proof: Deductive Reasoning in Human Thinking*. MIT Press.
Rips, Lance. 2002. Reasoning imperialism. In Renee Elio (ed.), *Common Sense, Reasoning, and Rationality*, 215–235. Oxford University Press.
Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5. 1–69.
van Rooij, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26(6). 727–763.
Russell, Stuart & Peter Norvig. 2010. *Artificial Intelligence: A Modern Approach*. Prentice Hall.
Sloman, Steven A. 2005. *Causal Models: How We Think About the World and its Alternatives*. Oxford University Press.
Spirtes, Peter, Clark Glymour & Richard Scheines. 1993. *Causation, prediction, and search*. MIT press.
Stone, Matthew. 1994. The reference argument of epistemic *must*. In *First International Workshop in Computational Semantics (IWCS 1)*, 181–190. http://www.cs.rutgers.edu/~mdstone/pubs/iwcs94.pdf.
Swanson, Eric. 2006. Interactions With Context. Ph.D. thesis, MIT.
Teigen, Karl. 1988. When are low-probability events judged to be 'probable'? Effects of outcome-set characteristics on verbal probability estimates. *Acta Psychologica* 68. 157–174.
Tenenbaum, Joshua B. & Tom L. Griffiths. 2001. Generalization, similarity, and

Bayesian inference. *Behavioral and Brain Sciences* 24(4). 629–640.
Tenenbaum, Joshua B., Tom L. Griffiths & Charles Kemp. 2006. Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences* 10(7). 309–318.
Tenenbaum, Joshua B., Charles Kemp, Tom L. Griffiths & Noah D. Goodman. 2011. How to grow a mind: Statistics, structure, and abstraction. *Science* 331(6022). 1279–1285.
Windschitl, Paul D. & Gary L. Wells. 1998. The alternative-outcomes effect. *Journal of Personality and Social Psychology* 75(6). 1411–1423.
Yalcin, Seth. 2005. A puzzle about epistemic modals. *MIT Working Papers in Linguistics* 51. 231–272.
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 A. Egan & B. Weatherson (eds.), *Epistemic Modality*, 295–332. Oxford University Press.

Daniel Lassiter
Department of Linguistics
Stanford University
450 Serra Mall, building 460
Stanford, CA 94305, USA
danlassiter@stanford.edu