

Trivializing Informational Consequence

Paolo Santorio ◦ University of Maryland, College Park

May 17, 2020

1 Overview

Recent work on epistemic modality appeals to nonclassical notions of logical consequence. On the classical conception (see e.g. Kaplan 1989a), logical consequence for natural language tracks preservation of truth (of a sentence, at a context). Many theorists have argued that this notion of consequence is inadequate for epistemic modal sentences and conditionals, like (1) and (2).

- (1) Frida might roll the die.
- (2) If Frida rolled the die, it came up even.

The details vary, but the central idea is that consequence for epistemic language should track instead a notion of preservation of support by an information state. On the resulting view, a conclusion B follows from a set of premises A_1, \dots, A_n just in case all bodies of information that support A_1, \dots, A_n also support B . This view of consequence naturally dovetails with a non-truth-conditional semantics, on which epistemic modal claims don't express propositions and are not true or false.

This paper investigates the link between informational consequence and credence. I first suggest a natural and seemingly harmless constraint concerning this link: informational consequence is certainty preserving. I.e., on any rational credence distribution, when the premises of an informational inferences have credence 1, the conclusion also has credence 1. This constraint has never been explicitly defended (though Kolodny and MacFarlane 2010 come close to doing so), but it dovetails with the widespread view that informational consequence tracks preservation of acceptance. It also allows us to make sense of intuitive judgments about informational inferences, including McGee-style alleged counterexamples to Modus Ponens.

After this setup, I show that, unfortunately, the certainty-preserving constraint leads to triviality. In particular, we can show that the following three claims are incompatible: (i) informational consequence is extensionally distinct from classical consequence (in particular, some inferences that are informationally valid are classically invalid); (ii) informational inferences are

certainty-preserving; (iii) credences are subject to (a subset of) classical Bayesian constraints. I also show that the result can be generalized to a case where we replace (ii) with the assumption that informational inferences preserves degree of credence n , for some n .

This result differs from all standard triviality results in the literature: it does not rely on probabilistic constraints about specific types of sentences (like, e.g., the constraint that the probability of $A > C$ is 1 whenever the probability of C is 1). Besides standard Bayesian tenets, the only assumptions are that informational consequence is distinct from classical consequence, and that informational consequence preserves certainty. At the same time, my proof provides a blueprint for generating triviality results for particular types of sentences, such as indicative conditionals. This suggests that informational inferences are a key ingredient in generating a certain family of triviality results, including classical results about conditionals.

The foregoing could be used as an argument against informational consequence, but this is not my goal. The upshot of this paper is that informational consequence requires a nonstandard account of credence and credal update. I take up this task elsewhere, but in the last section I offer some pointers.

I proceed as follows. §2 sets up a simple version of the informational view, illustrating some of its advantages. §§3–§4 spell out the question of the link between credence and consequence and motivate the certainty preservation constraint. I present the new triviality result in §5, and discuss the options for informational theorists in §6. A terminological note: throughout the paper I use ‘inference’ as a synonym of ‘entailment’, i.e. to characterize pairs of a set of premises and a conclusion.

2 Setup: informational consequences

This paper focuses on logical consequence, but it is impossible to illustrate how consequence applies to natural language without introducing a semantics. So I start by setting up a simple semantics for epistemic modals and conditionals, and defining classical and informational consequence for it. I choose a semantics that closely mimics the semantics in Yalcin 2007.

2.1 Semantics for epistemic modals

I use an interpretation function (represented as ‘ $\llbracket \cdot \rrbracket$ ’) to map expressions to their semantic values. As usual, this mapping is relativized to an n -tuple of parameters (a point of evaluation). I take points of evaluation to be a pair of a world and an information state $\langle w, i \rangle$. Hence the general form of a semantic clause is:

$$\llbracket A \rrbracket^{w,i} = \text{semantic value of } A \text{ relative to } \langle w, i \rangle$$

The world and information state parameters are exploited selectively by different fragments of the language. Nonmodal sentences are sensitive to the world parameter, but not the information state parameter. Here is an example of a semantic clause for a simple, nonmodal sentence:

$$(3) \quad \llbracket \text{It is raining} \rrbracket^{w,i} = \text{true iff it is raining at } w$$

Conversely, modal operators display sensitivity to the information state parameter, but not to the world parameter. In particular, necessity and possibility modals are analyzed as quantifiers over worlds in the information state picked out by i :

$$(4) \quad \llbracket \Box A \rrbracket^{w,i} = \text{true iff } \forall w' \in i : \llbracket A \rrbracket^{w',i} = \text{true}$$

$$(5) \quad \llbracket \Diamond A \rrbracket^{w,i} = \text{true iff } \exists w' \in i : \llbracket A \rrbracket^{w',i} = \text{true}$$

This semantics can be naturally extended to indicative conditionals. Here I use a simple variant of the semantics in Gillies (2004, 2009). Start by defining a notion of update of an information state:

Update of i with A

$$i + A = i \cap \{w : \llbracket A \rrbracket^{w,i} \text{ is true}\}$$

Conditional antecedents are used to update the information state in the index; conditional consequents are evaluated at the updated information state. Using the traditional corner ' $>$ ' to represent the conditional, here is the relevant clause:

$$(6) \quad \llbracket A > B \rrbracket^{w,i} = \text{true iff } \forall w' \in i + A : \llbracket B \rrbracket^{w',i+A} = \text{true}$$

2.2 Defining consequence

The compositional semantics I just outlined can be made fully compatible with a conservative picture of content and consequence. In particular, we may define a notion of truth at a context in the style of Kaplan (1989a, 1989b):

$$(7) \quad A \text{ is } \mathbf{true at } c \text{ iff } \llbracket A \rrbracket^{w_c, i_c} = \text{true}$$

This allows us to assign classical truth values to utterances of epistemic sentences. (Of course, the definition in (7) requires a metasemantic assumption: each context determines an information state that is relevant for evaluating an utterance. This assumption is widely disputed.¹)

Proponents of informational consequence take an alternative path. They refrain from defining a standard notion of truth. Instead, they define a notion of support by an information state. Intuitively, a sentence is supported by an information state just in case an agent who is in that state accepts the sentence. Formally:

¹See, among many, Egan et al. 2005, MacFarlane 2011, 2014.

i **supports** A ($i \vDash A$) iff, for all $w \in i$, $\llbracket A \rrbracket^{w,i} = 1$.

For the case of nonmodal sentences, A being supported by i simply reduces to A being true at all worlds in i . But for modal sentences, i supports A just in case i satisfies a kind of global condition, not reducible to properties of each individual worlds. For example, i supports $\Diamond A$ just in case i contains some world where A is true.

These two views are naturally paired with different notions of consequence. On the classical truth conditional picture, consequence may be defined in the standard way, as preservation of truth at a point of evaluation. In particular, one notion of consequence that is of particular relevance is preservation of truth at a *proper* point, i.e. a pair $\langle w, i \rangle$ such that $w \in i$.² This is the notion that, throughout the paper, I will refer to as ‘classical consequence’.

Classical logical consequence.

$A_1, \dots, A_n \vDash_C B$ iff for all $\langle w, i \rangle$ such that $w \in i$ and $\llbracket A_1 \rrbracket^{w,i} = 1, \dots, \llbracket A_n \rrbracket^{w,i} = 1, \llbracket B \rrbracket^{w,i} = 1$.

Conversely, informational consequence is defined as preservation of support.

Informational consequence.

$A_1, \dots, A_n \vDash_I B$ iff, for all i s.t. i supports A_1, \dots, A_n , i supports B .

Informational consequence is, at least in a sense, a special case of classical consequence. We could define it from classical consequence by restricting consideration not just to proper points of evaluation, but also to points of evaluation where the premises of an argument are true at all worlds in the information state. We get:

Classical* logical consequence.

$A_1, \dots, A_n \vDash_{C^*} B$ iff for all $\langle w, i \rangle$ such that $w \in i$ and such that, for all $w' \in i$, $\llbracket A_1 \rrbracket^{w',i} = 1, \dots, \llbracket A_n \rrbracket^{w',i} = 1, \llbracket B \rrbracket^{w,i} = 1$.

The reader can check that classical* consequence is a notational variant of informational consequence.

The fact that informational consequence is a special case of classical consequence means that informational consequence is strictly weaker than classical consequence, in the sense that it is more permissive. On the one hand, all classically valid rules of inferences are also informationally valid:

Fact. For all A_1, \dots, A_n, B : if $A_1, \dots, A_n \vDash_C B$, then $A_1, \dots, A_n \vDash_I B$.

On the other hand, some extra rules of inferences, which are not classically valid, are informationally valid. One example of a rule of inference that is classically invalid but informationally valid is Łukasiewicz’s principle:

²Why is this notion ‘particularly relevant’? Differently from improper points, proper points represent genuine epistemic predicaments, i.e. pairs of a world and an informational state such that, for all a subject knows, they might be located at. The resulting notion of consequence is equivalent to what Yalcin (2007) calls ‘diagonal consequence’.

Lukasiewicz's principle. (ŁP) $\neg A \vDash \neg \diamond A$ (Equivalently³: $A \vDash \Box A$)

Lukasiewicz's principle is at the center of arguments for the informational view. For example, Yalcin 2007 argues for the informational view by arguing that sentences like (8) are semantically contradictory, and that this is captured by informational and not classical consequence.

(8) #It's not raining and it might be raining.

Also with regard to conditionals, informational consequence vindicates plausible inferences that are not classically valid. In particular, consider Modus Ponens:

Modus Ponens. $A > B, A \vDash B$

A classical argument due to Gibbard (1981) shows that Modus Ponens, in its unrestricted form, is classically incompatible with other plausible logical principles. One upshot is that, on several classical semantics for conditionals, such as Kratzer's (1986, 2012), Modus Ponens holds only in restricted form, i.e. only when A and B involve no modality and no conditionals. (See Khoo 2013 for discussion.)⁴ Conversely, informational consequence allows us to vindicate the full strength of Modus Ponens (see Bledin 2015).

A qualification: the claim that informational consequence validates strictly more inferences than classical consequence doesn't include what we may call 'meta-rules', i.e. rules that allow us to infer that a conclusion follows from a set of premises on the basis of the fact that other entailments hold. On the contrary: classical meta-rules fail in informational consequence. I discuss the distinction between rules of inference and meta-rules in a footnote.⁵

⁴More precisely, Gibbard presents a so-called collapse result. He shows that, given classical logic, three plausible principles of conditional logic lead to an implausible conclusion. The principles in question are:

Upper bound. If $A \vDash B$, then $\vDash A > B$
Centering. $A > B \vDash A \supset B$
Exportation. $(A \wedge B) > C \vDash A > (B > C)$

Both Centering and Exportation are empirically plausible; moreover, Centering is what guarantees the validity of Modus Ponens for natural language conditionals (see, among others, Gillies 2004 and 2009 for presentation of relevant data). Yet Gibbard shows that, given the three principles above, we can prove that an indicative conditional $\lceil A > B \rceil$ is logically equivalent to the corresponding material conditional $\lceil A \supset B \rceil$. Here is a slightly updated version of Gibbard's proof. (For different versions of the proof, see McGee 1985, Gillies 2009.) One direction of the biconditional is provided by Centering; here is a proof of the other direction.

- (i) $A \supset B$
- (ii) $(A \wedge B) > B$ (by Upper Bound)
- (iii) $((A \supset B) \wedge A) > B$ ((ii), propositional logic)
- (iv) $(A \supset B) > (A > B)$ ((iii), Exportation)
- (v) $(A \supset B) \supset (A > B)$ ((iv), Centering)
- (vi) $A > B$ ((i), (v), propositional logic)

⁵A rule of inference has the form:

3 The question: informational consequence and credence

Several theorists take informational consequence to be the notion of consequence that correctly captures the logic of epistemic discourse. This attitude is pervasive in the dynamic semantics literature (for the *locus classicus*, see Veltman 1996)⁶. More recently, it has become widespread among philosophers. In this vein, Gillies 2009 has claimed that "entailment ought not be flat-footed [preservation of truth at a world]" (p. 343)⁷; Yalcin 2012 has suggested that a classical account of consequence may not be "adequate for modeling natural language" (p. 1011); Bledin 2015 has argued that logic—insofar as it applies to natural language—is "fundamentally concerned not with the preservation of truth but rather with the preservation of structural properties of ... bodies of information" (p. 64).

In this paper, I focus on the link between informational consequence and credence. If informational consequence captures the logic of epistemic discourse, we should expect that informational consequence somehow places constraints on credences in epistemic sentences. I take this question to split into a normative and a descriptive subquestion. Respectively:

- (i) What rational constraints, if any, link informational consequence and credence?
- (ii) What constraints do subjects's actual credences in epistemic claims actually conform to (at least, by and large)?

In the next section, I suggest a constraint that targets primarily the normative question. The constraint specifies how credences of *rational* subjects work. It also appears that credences of actual subjects, by and large, abide by the constraint, though I won't need this assumption for my argument.⁸

$$A_1, \dots, A_n \vDash_i B$$

I.e., a rule of inference allows the derivation of a sentence in the object language from a set of sentences in the object language. Conversely, a meta-rule has the form:

$$\text{If } A_1, \dots, A_n \vDash_i B \text{ and } \dots \text{ and } C_1, \dots, C_n \vDash_i \theta, \text{ then } \sigma_1, \dots, \sigma_n \vDash_i \omega$$

I.e., a meta-rule allows us to derive that a certain rule of inference holds, given that other rules of inference hold. Informational consequence does *not* validate more meta-rules than classical consequence. In fact some meta-rules, like *reductio* and proof by cases, fail on informational consequence for the fragment of the language that involves epistemic modality. (See for example Willer 2012 and Bledin 2015 for informational semantics that restrict the use of proof by cases.) This is not surprising: if we add some rules of inference to a logic, there is a risk that meta-rules will be invalidated, since the antecedent of meta-rules ends up ranging over some extra cases.

⁶More precisely, Veltman presents three notions of validity in his 1996; informational consequence corresponds to what he labels 'validity₃'. All three of Veltman's notions, however, track a kind of preservation of support (as opposed to preservation of truth).

⁷Gillies doesn't endorse informational consequence as I present it, but rather opts for a more dynamic variant of it (so-called update-to-test consequence, in the terminology of Veltman 1996). But, for current purposes, it seems fair to place him in the same broad camp as theorists like Yalcin and Bledin.

⁸The existence of normative constraints linking consequence and credence is controversial. (For

4 Informational consequence and preservation of certainty

In this section, I outline and motivate a natural view: informational consequence is certainty preserving, in the following sense: any rational subject that is certain of the premises of an informational inference is also certain of the conclusion. This view has never been formulated explicitly (though see Kolodny and MacFarlane 2010 for remarks going in this direction), but it seems natural and desirable. It accommodates several observations in the literature, and it allows us to make sense of some difficult examples, including alleged McGee-style counterexamples to Modus Ponens. To motivate it, I start from two examples that raise *prima facie* challenges for informational consequence.

4.1 Credence and Łukasiewicz's principle

Recall that informational consequence validates:

Łukasiewicz's principle. (ŁP) $\neg A \vDash \neg \Diamond A$ (Also: $A \vDash \Box A$)

Moritz Schulz (2010) has claimed that Łukasiewicz's principle doesn't track sound probabilistic reasoning. (For a similar point concerning closure and informational consequence, see also Bledin and Lando 2018.) To make his point, Schulz simply points out that one may reasonably assign high credence to (9)–a, while assigning "very low credence (or even credence 0)" to (9)–b.

- (9) a. They are at home.
 b. They must be at home.

Schulz explicitly assumes that a plausible notion of consequence should preserve probability in one-premise inferences. Formally:

Single-Premise Closure (SPC).

If $A \vDash B$ and $Pr(A) = t$, $Pr(B) \geq t$.

SPC seems to be a basic closure principle that governs all interaction between logic and probability. Since Łukasiewicz's principle violates SPC, Schulz rules it out as invalid.

This clearly violates a reasonable constraint on logical consequence: If a rational and logically omniscient subject's credence function P is such that $P(A) = t$, and $A \vDash B$, then $P(B) \geq t \dots$ *Since this seems to be one of the core features of how logical consequence relates to rational reasoning, we should not accept informational consequence as an account of the logic of epistemic modals.* (Schulz 2010)

the debate about the normative import of logical consequence and how it relates to attitudes, see e.g. Christensen 2004, MacFarlane 2004, Field 2015 for the view that logic imposes normative constraints on attitudes; and see Harman 1984 for the denial of that claim. See also Kolodny 2007, 2008 for discussion.) Here let me just notice that all Bayesian accounts appear to be committed to the existence of constraints of this sort.

To my knowledge, Schulz’s point has received no discussion in the literature. In part, this may be due to skepticism towards the idea that *might*- and *must*-claims are appropriate objects of credence.⁹ But Schulz’s point is fully general. In particular, it also applies also to conditionals, for which the ‘no probability’ claim seems harder to make.

4.2 Credence and Modus Ponens

The next case I want to discuss involves Modus Ponens. Recall from §2 that Modus Ponens is informationally valid but not (not in general, at least) classically valid. Below I argue that Modus Ponens, on a par with Łukasiewicz’s principle, fails as a general constraint on credence. My examples are reminiscent in obvious ways of McGee’s well-known examples about Modus Ponens, though I focus specifically on credal judgments.

Start from a simple formulation of Modus Ponens:

Modus Ponens. $A > B, A \vDash B$

Now, assume the following generalization of Single-Premise Closure:

Multi-Premise Closure (MPC).

If $A_1, \dots, A_n \vDash B$ and $Pr(A_1) = t_1, \dots, Pr(A_n) = t_n$, then

$Pr(B) \geq (1 - ((1 - t_1) + \dots + (1 - t_n)))$.

One intuitive gloss for MPC is the following: a subject’s degree of confidence that the conclusion of an inference is false should not exceed the sum of their degrees of confidence that each of the premises is false. This kind of closure principle is widely adopted (see e.g. Adams 1975 and Field 2015).

From MPC and Modus Ponens the following constraint follows:

Probabilistic Modus Ponens. (PMP)

If $Pr(A > B) = 1$, then $Pr(A) \leq Pr(B)$.

PMP says that, if a conditional has probability 1, then the probability of the antecedent cannot exceed the probability of the consequent. It is a natural probabilistic generalization of Modus Ponens.¹⁰

Even if we reject MPC, we can derive PMP from its single-premise counterpart SPC, plus other minimal constraints on credence. In particular, consider:

Conjunction Lower Bound. (CLB)

If $Pr(A) = 1$, $Pr(A \wedge B) = Pr(B)$.

⁹As a sociological note: I myself have repeatedly encountered this attitude in conversation.

¹⁰PMP is a special case of a more general principle, which also follows from standard MP and MCP:

Generalized Probabilistic Modus Ponens.

Let $Pr(\text{if } A, B) = 1 - d$. Then $Pr(A) - d \leq Pr(B)$.

CLB says that, if one has credence 1 in a conjunct, then one's credence in a conjunction is equal to the credence in the other conjunct. CLB is a fairly weak constraint, which is validated by classical probability as well as several nonstandard probability theories.¹¹ But, together with SPC, it is sufficient to entail PMP.

Now, consider probabilistic judgments about conditionals in a scenario that is reminiscent of the scenarios in McGee's (1985). Suppose that Sarah has tossed a fair six-sided die. For convenience, name the die 'Die'. You have no information about the outcome of the toss. Now consider:

(C) If Die did not land on two or four, then it landed on six.

I ask you to pause and consider what level of credence you assign to (C)—whether low, middling (roughly, 'fifty-fifty'), or high.

In informal polls, most subjects answer 'low'. (Several people also give the more precise answer '1/4'.) Very few people, if any, go for 'middling' or 'high'. Now consider:

(P2) Die landed even.

(P1) If Die landed even, then, if it didn't land on two or on four, it landed on six.

In informal polls, (P2) and (P1) are assigned, respectively, middling and high credence.¹²

The table below summarizes the judgments about the die scenario. I model the judgment of certainty concerning (P1) as credence 1, though this assumption can be weakened without harm.¹³

(P1)	if even, (if not (two or four), six)	certain (=1)
(P2)	even	middling ($\approx .5$)
(C)	if not (two or four), six	low ($\approx .25$)

These judgments provide a counterexample to PMP. On the assumption that MPC, or SPC plus CLB, are not in question (see below), this suggests that

¹¹For example, it is validated by the nonclassical probability theories based on Kleene logics that are discussed in Williams 2016.

¹²For convenience, I will talk of credences as being modeled by probability functions, but I only need the assumption that credences conform to MPC (or SPC plus CLB). Also, I don't need the assumption that subjects assign precise numerical values to claims. All I need is that subjects make coarse-grained distinctions between 'high', 'middling', and 'low' degrees of belief. Also, a bookkeeping note: for the moment, I assume that probabilities attach directly to statements (construed as sentences as uttered at a context).

¹³Suppose we assign (P1) credence $1 - \epsilon$, for some tiny ϵ . On this understanding, PMP simply doesn't concern (P1). However, as long as the difference in credences between (P2) and (C) is greater than ϵ , the same statements violate the generalization of PMP given in footnote 10. For simplicity, from now on I simply assume that the rational credence in (P1) is 1.

Modus Ponens fails as a constraint on credence in natural language.¹⁴

As an aside: the foregoing sheds light on McGee’s classical discussion of Modus Ponens (1985). McGee considers examples just like our (P1)-(C), against the backdrop of scenarios like the following:

Sarah tossed Die, a fair six-sided die. You caught a brief glimpse of the face that came up on top. You are confident, though not certain, that it landed on 2. In any case, you are very confident that you saw few dots on the face that came up.

In McGee’s words (1985, p. 462), in a scenario like this you have “good grounds for believing the premises [i.e. (P1) and (P2)]”, but you are “not justified in accepting the conclusion”. McGee concludes that:

sometimes the conclusion of an application of Modus Ponens is something we do not believe and should not believe, even though the premises are propositions we believe very properly (1985, p. 462).

I suggest that McGee’s claim here is best interpreted as a claim about credal drop. What McGee observes is that, in some scenarios, a subject may rationally have high credence in the premises of a Modus Ponens argument and low credence in the conclusion, in a way that is incompatible with validity. On this interpretation, McGee’s claim is correct.

4.3 Preserving certainty

I have pointed out that there are rational credence distributions that are incompatible with the credal validity of some informational inferences (assuming minimal closure constraints on credence). This suggests that informational consequence does not track rational constraints on credence in general.

At the same time, as Gillies 2004 emphasizes, inferences like Łukasiewicz’s principle and Modus Ponens sound invariably valid, even in McGee-style cases, in contexts in which the premises are accepted. For example, if a subject accepts (P1) and (P2), they also appear to be committed to accepting (C).

(P1) If Die landed even, then, if it didn’t land on two or on four, it landed on six.

(P2) Die landed even.

¹⁴A line of objection to the foregoing (represented e.g. in Stojnic 2017) is that the argument doesn’t really represent an instance of Modus Ponens, since the unembedded occurrence of the conditional *if not (two or four), six* exploits a different domain of quantification from the embedded one. Much can be said to rebut this objection. For current purposes, I can simply sidestep it. I can agree with the objector that MP proper is not threatened by (P1)-(C). What matters for me is not whether Modus Ponens, in whatever way it should be defined, is valid or invalid. My goal is establishing that some inference patterns that are informationally valid fail qua constraints on probabilistic attitudes. This point is independent of what we choose to call ‘Modus Ponens’. Dub the principle exemplified by (P1)-(C) ‘MP*’. MP* illustrates the failure I’m interested in.

(C) If Die did not land on two or four, then it landed on six.

In fact, just an observation of this sort is at the basis of Stalnaker's notion of *reasonable inference*, which is an ancestor of informational consequence.¹⁵ This suggests a natural constraint linking informational consequence and acceptance:

Acceptance Preservation

If $A_1, \dots, A_n \vDash_I B$, then for all i modeling the information states of rational subjects: either one of A_1, \dots, A_n is not supported by i , or B is supported by i .

Following Stalnaker (1984, 2002), I use 'acceptance' to refer to a broad category of mental states that includes but goes beyond belief. On a first pass, acceptance is the broadest possible kind of doxastic attitude. Accepting a proposition consists in taking it as true for some purposes or other.

Now, one notable special case of acceptance is certainty, which in a credal framework is equivalent to credence 1. So I take the following to be a special case of **Acceptance Preservation**:

Certainty Preservation

If $A_1, \dots, A_n \vDash_I B$, then, for all Pr modeling rational credence:
if $Pr(A_1) = 1, \dots, Pr(A_n) = 1$, then $Pr(B) = 1$

Certainty Preservation seems obvious. Consider again to the die example: if you assign credence 1 to (P1) and (P2), it seems obvious that you should also assign credence 1 to (C).

Notice that **Certainty Preservation** (as well as **Acceptance Preservation**) is a synchronic constraints. I.e., it states consistency facts for attitudes at a given time. I make no assumptions about dynamic constraints linking consequence and credence.

Some theorists might suggest that acceptance should not be linked to full credence, but rather to a high enough degree of credence, or credence that exceeds a certain threshold. I am skeptical of this identification, just because of the data discussed in this section. Łukasiewicz's Principle and Modus Ponens do *not* appear to be valid when the premises are assigned any degree of credence that falls short of 1. In any case, switching to a threshold-type view will not block the main argument. In §5, I show how the triviality result can be generalized to a case where we take informational consequence to preserve credence above t , for some value of t .

¹⁵Here is Stalnaker's definition:

[A]n inference from a sequence of assertions or suppositions (the premisses) to a conclusion is a reasonable inference just in case in every context in which the premisses can be asserted or supposed, it is impossible for anyone to accept the premisses without committing himself to the conclusion. (1975, p. 271)

See Cariani 2020 for a discussion of Stalnaker's notion.

4.4 Side note: Credence Preservation

Certainty Preservation is all that I need to run my triviality argument. No further assumptions about the link between logic and credence are needed. But before moving on, let me suggest that we should also accept a further principle. This point is independent from the main line of argument of the paper, but it helps fit **Certainty Preservation** in a broader picture of the relation between credence and logic for epistemic discourse.

The second principle is simply that classical consequence is the notion of consequence that tracks preservation of degree of credence.¹⁶

Credence Preservation

If $A_1, \dots, A_n \vDash_C B$, then for all Pr modeling rational credence: if $Pr(A_1) = t_1, \dots, Pr(A_n) = t_n$, then $Pr(B) \geq (1 - ((1 - t_1) + \dots + (1 - t_n)))$.

Credence Preservation says, in effect, that classical consequence is the notion of consequence regulating credence. **Credence Preservation** can coexist with **Certainty Preservation**. If we accept both, we assign a role to both classical and informational consequence. The two track different constraints on attitudes. Classical consequence tracks constraints on probabilistic reasoning in general. Informational consequence tracks constraints on reasoning from premises that are certain. Among other things, this view seems to account nicely for the evidence presented on both sides of the debate about informational consequence. Łukasiewicz's Principle and Modus Ponens are probabilistically invalid in general, but valid when taken to be inferences from the premises that are certain.

This split view has not been defended explicitly anywhere in the literature. But Kolodny and MacFarlane (2010) make a suggestion in this direction in their discussion of Modus Ponens. They argue that Modus Ponens is invalid, on the basis of examples involving deontically modalized consequents. But they point out that, even on their view, Modus Ponens is *quasi-valid*, where an argument is quasi-valid iff the argument that we obtain from it by prefacing all the premises with epistemic necessity modals is valid. Quasi-validity is an object language counterpart of the view that Modus Ponens and other informational inferences govern reasoning from accepted premises.

5 Informational triviality

The previous section suggests a natural constraint about the link between informational consequence and credence: on any rational credence distribution,

¹⁶A bridge principle of this sort has been recently defended by Field 2015. Here is a statement that more closely mirrors Field's own statement:

If $A_1, \dots, A_n \vDash_C B$, then a subject's credence function Cr_S is such that: $Cr(B) \geq \sum_i Cr(A_i) - n + 1$.

when the premises of an informational inferences have credence 1, the conclusion also has credence 1. Unfortunately, this view is incompatible with broadly Bayesian tenets. In this section, I prove a general result, which is in the broad family of triviality results about conditionals and modals.¹⁷

Differently from other triviality theorems, this result does not rely on assumptions about the semantics of particular expressions, like conditionals or epistemic modals. The only starting assumptions are (i) that there is a notion of consequence that is weaker than classical consequence (in the sense that more inferences turn out to be valid), and (ii) that this notion of consequence preserves credence 1. The upshot is that we cannot have all of the following: (i) a notion of consequence that is extensionally different from classical consequence; (ii) **Certainty Preservation**; (iii) some classical assumptions about credence and credal update, which I introduce below under the labels of Identity, Bound, Closure, and Plenitude.

This leaves the informational theorist with a number of options, which I examine in detail in §6. But the general upshot is that endorsing informational consequence requires adopting a nonstandard view about credence—provided that we want to assign credence to epistemic sentences in the first place.

5.1 Triviality for informational consequence

Start from the assumption that we have an interpreted language \mathcal{L} , for which we define a classical and an informational notion of consequence $\vDash_{C_{\mathcal{L}}}$ and $\vDash_{I_{\mathcal{L}}}$. Following Bradley 2000, I will speak of a language being trivial; I will understand the notion as follows:

\mathcal{L} is **trivial** iff:
 for any A_1, \dots, A_n, B : if $A_1, \dots, A_n \vDash_{I_{\mathcal{L}}} B$, then $A_1, \dots, A_n \vDash_{C_{\mathcal{L}}} B$.
 (\mathcal{L} is **nontrivial** otherwise.)

I.e., a language is trivial iff every informational inference is also a classical inference. Hence a trivial language makes no room for a notion of consequence that is distinct from classical consequence.¹⁸

Assume now that we have a language \mathcal{L} , for which we define two notions of consequence $\vDash_{C_{\mathcal{L}}}$ and $\vDash_{I_{\mathcal{L}}}$. To start, I assume **Certainty Preservation** from §4, repeated below.

Certainty Preservation

If $A_1, \dots, A_n \vDash_I B$, then, for all Pr modeling rational credence:
 if $Pr(A_1) = 1, \dots, Pr(A_n) = 1$, then $Pr(B) = 1$

¹⁷A very incomplete list of helpful references on triviality may include Lewis 1976, Hájek and Hall 1994, Bradley 2000, Bradley 2007, Russell and Hawthorne 2016, Goldstein forthcoming; see Khoo and Santorio 2018 for an introductory overview.

¹⁸This notion of triviality is obviously different from other notions of triviality appearing in the literature, but this is irrelevant. What matters is that we get collapse of one notion of consequence onto another.

Then, I appeal to four additional assumptions about credence and conditional credence. The first two are entailed by classical Bayesianism, the other two are *prima facie* very plausible.¹⁹

1. Identity. The first assumption is very simple: one’s credence in A, conditional on A, should be 1.

Identity. $Pr_A(A) = 1$

For current purposes, I understand the conditional probability $Pr(\bullet | A)$ (which I denote interchangeably with ‘ $Pr_A(\bullet)$ ’) as the probability function that we obtain by rationally updating $Pr(\bullet)$. For the purposes of the proof I don’t need to assume that the conditional probability $Pr(\bullet | A)$ is defined via the Ratio formula, though all that I say is compatible with this assumption.

2. Bound. The second assumption is a constraint linking the conditional probability of B given A, and unconditional probabilities of A and B.

Bound. If $Pr(B | A) = 1$, then $Pr(B) \geq Pr(A)$

Bound is entailed by standard Bayesianism. Moreover, it can be validated even in nonclassical construals of probability, provided that we have two assumptions. One is CLB from §4, i.e. the principle that the probability of a conjunction is a lower bound on the probability of the conjuncts. The other is the ratio construal of conditional probability:

Ratio. $Pr(B | A) = \frac{Pr(A \wedge B)}{Pr(A)}$ (with $Pr(A) > 0$)

CLB and **Ratio** are sufficient to entail **Bound**.

3. Closure. The third assumption is that the class of rational credence functions is closed under conditionalization, in the following sense: if we have a rational credence function Pr , then for any A, the credence function Pr_A that we obtain by conditionalizing Pr on A is also rational.

Closure. For all A: If $Pr(\bullet)$ models a rational credence function, then $Pr(\bullet | A)$ also models a rational credence function.

Closure is appealed to in triviality proofs without qualification (for example by Lewis 1976). According to standard Bayesians, rational agents update by conditionalizing on the total evidence that they learn. **Closure** follows from this if we assume that any proposition can be an agent’s total evidence. **Closure** has recently come under attack from a few directions (see e.g. Author/s 2019a). I assume it for the purposes of the proof; as I point out in §6, one of the possible morals to draw is just that endorsing informational consequence requires jettisoning **Closure**.

¹⁹The labels ‘Identity’ and ‘Bound’ are borrowed from Author/s 2019b.

4. Plenitude. The fourth assumption is a principle linking credence and classical consequence. It says that, if A_1, \dots, A_n do not classically entail B , there is a rational credence function that assigns higher credence to the conjunction of A_1, \dots, A_n than to B .

Plenitude. If $A_1, \dots, A_n \not\vdash_{C_{\mathcal{L}}} B$, then for some Pr modeling rational credence, $Pr(A_1 \wedge \dots \wedge A_n) > Pr(B)$.

Plenitude is the least familiar of the four assumptions, and at first sight the one that might look most questionable. I defend it below.

I will also need an elementary principle about entailment, stating that if a set of premises entails a conclusion, then the conjunction of those premises entails that conclusion:

Conjoined Premises. If $A_1, \dots, A_n \vDash_{C_{\mathcal{L}}} B$, then $A_1 \wedge \dots \wedge A_n \vDash_{C_{\mathcal{L}}} B$.

I take this principle to be safe enough that I won't give it any further discussion.

Given these assumptions, we prove the following result:

Theorem: informational Triviality.

Given **Certainty, Bound, Closure, Plenitude**: for any \mathcal{L} , \mathcal{L} is trivial.

PROOF. For reductio, assume that, for some A_1, \dots, A_n and B , $A_1, \dots, A_n \vDash_I B$ and $A_1, \dots, A_n \not\vdash_C B$. Via **Plenitude**, we have that there is a rational credence function Pr such that $Pr(A_1 \wedge \dots \wedge A_n) > Pr(B)$. Via **Closure**, we know that $Pr_{A_1 \wedge \dots \wedge A_n}$ is also a rational credence function, and via **Identity** we know that $Pr_{A_1 \wedge \dots \wedge A_n}(A_1 \wedge \dots \wedge A_n) = 1$. Contraposing on **Bound**, since $Pr(A_1 \wedge \dots \wedge A_n) > Pr(B)$ we know that $Pr_{A_1 \wedge \dots \wedge A_n}(B) < 1$. But via **Certainty Preservation** (and using **Conjoined Premises**), since $Pr_{A_1 \wedge \dots \wedge A_n}(A_1 \wedge \dots \wedge A_n) = 1$, we also have $Pr_{A_1 \wedge \dots \wedge A_n}(B) = 1$. Contradiction.

5.2 Plenitude and 'localized' triviality

We proved that we cannot have all three of: informational consequence; a plausible constraint on credence stating that informational consequence preserves credence 1; and some fairly standard assumptions about credence (some of them part of Bayesianism) and credal update. In the next section, I chart the options for the informational theorist in the face of this predicament.

Before this, let me say something in defense of my last assumption, i.e. **Plenitude**. Stated in English, it says that, if a set of premises A_1, \dots, A_n doesn't classically entail a conclusion B , there is a rational credence function that assigns higher credence to the conjunction of A_1, \dots, A_n than to B .²⁰

²⁰In stating **Plenitude** in this way, I am assuming that conjunction behaves in a classical way. We could complicate the statement of **Plenitude** to abstract from this assumption.

This assumption might seem unfamiliar (though it's worth noticing that, in the proof of his well-known triviality result, Bradley 2000 simply assumes the analog of **Plenitude** without argument). Let me list a few considerations in its defense.

First, the examples about Łukasiewicz's Principle and Modus Ponens discussed in §4 provide inductive support for **Plenitude**. We observed precisely that there are rational credence functions that assign higher credence to the conjunction of the premises of an informational inference than to the conclusion. I invite the reader to check other examples for other informational inferences.

Second, while specifying a model of credence for informational content goes beyond the purposes of this paper, it's easy to see how a plausible model of this sort will validate **Plenitude**. For concreteness, suppose that we use a Yalcin-style model of content, and we assign probabilities to sentences by assigning probabilities to points of evaluation, i.e. pairs $\langle w, i \rangle$ of a world and an information state. We know that, if an inference is classically invalid, there is a pair $\langle w, i \rangle$ at which all the premises are true and the conclusion is false. Absent further constraints on priors, this guarantees that there is a credence function satisfying **Plenitude**.

Third, even if **Plenitude** in full generality were to fail, we could still get special cases of **Informational Triviality**. In particular, special versions of the theorem would hold for all informational inferences for which **Plenitude** holds. (Again, §4 shows precisely some examples of these inferences.) This is damaging enough. It would show that we cannot have all of the following: Łukasiewicz's principle (say) is informationally but not classically valid; Łukasiewicz's principle preserves certainty; and the package of classical assumptions about credence holds.

This third consideration also points to a further feature of **Informational Triviality**. The result can be used as a blueprint for generating a series of 'localized' triviality results about individual inferences. For concreteness, take Łukasiewicz's principle. Suppose we replace **Certainty Preservation** with the following:

$$\text{If } Pr(\neg A) = 1, \text{ then } Pr(\neg \diamond A) = 1$$

We can use the proof in §5.1 to show that we cannot have all the following: Łukasiewicz's Principle is informationally valid but classically invalid, Łukasiewicz's Principle preserves certainty; our assumptions (1)–(4). *Mutatis mutandis*, a similar strategy holds for all informational inferences.

5.3 Generalizing the result

Let me consider a way of resisting the result that we have reached in this section.²¹ I have assumed that **Certainty Preservation** is a special case of **Accep-**

²¹Thanks to an anonymous referee for pushing me to consider this option.

tance Preservation. But one may actually deny this. **Acceptance Preservation**, they may argue, does not require preservation of credence 1, but rather preservation of a degree of credence that exceeds a certain threshold t . So **Certainty Preservation** should be replaced by:

Threshold Preservation

If $A_1, \dots, A_n \vDash_I B$, then, for all Pr modeling rational credence:
 if $Pr(A_1) \geq t, \dots, Pr(A_n) \geq t$, then $Pr(B) \geq t$
 (with $t = 1 - \epsilon$, for some ϵ)

The result can be reproduced also starting from **Threshold Preservation**, if we assume modified versions of **Bound** and **Plenitude**.

Instead of **Bound**, we assume:

Threshold Bound.

If $Pr(B | A) \geq t$, then $Pr(A) - Pr(B) < \epsilon$ (with $t = 1 - \epsilon$)

Threshold Bound says that, if one's credence in B conditional on A is higher than a threshold $1 - \epsilon$, the maximum difference between one's credence in A and one's credence in B is ϵ . Like **Bound**, **Threshold Bound** is also entailed by standard Bayesianism.

Instead of **Plenitude**, we assume:

Threshold Plenitude. If $A_1, \dots, A_n \not\vDash_{C_L} B$, then for some Pr modeling rational credence, $Pr(A_1 \wedge \dots \wedge A_n) - Pr(B) > \epsilon$.

Threshold Plenitude is strictly stronger than **Plenitude**, but similar arguments in its support apply. In particular, notice that the examples discuss in §4 support also **Threshold Plenitude**.

At this point, it's easy to state the proof of **Informational Triviality** on the new assumptions.

PROOF. For reductio, assume that, for some A_1, \dots, A_n and B, $A_1, \dots, A_n \vDash_I B$ and $A_1, \dots, A_n \not\vDash_C B$. Via **Threshold Plenitude**, we have that there is a rational credence function Pr such that $Pr(A_1 \wedge \dots \wedge A_n) - Pr(B) > \epsilon$. Via **Closure**, we know that $Pr_{A_1 \wedge \dots \wedge A_n}$ is also a rational credence function, and via **Identity** we know that $Pr_{A_1 \wedge \dots \wedge A_n}(A_1 \wedge \dots \wedge A_n) = 1$. Contraposing on **Threshold Bound**, since $Pr(A_1 \wedge \dots \wedge A_n) - Pr(B) > \epsilon$, we know that $Pr_{A_1 \wedge \dots \wedge A_n}(B) < t$. But via **Threshold Preservation** (and using **Conjoined Premises**), since $Pr_{A_1 \wedge \dots \wedge A_n}(A_1 \wedge \dots \wedge A_n) \geq t$, we also have $Pr_{A_1 \wedge \dots \wedge A_n}(B) \geq t$. Contradiction.

A different way of denying the link between credence 1 and acceptance involves denying that epistemic sentences can be the object of credence in the first place. I examine this view in the next section.

6 Fitting probability into the informational view

In this section, I outline what I take to be the two main options for the informational theorist. The first involves denying that probability applies at all to epistemic sentences; the second involves moving to a nonclassical notion of probability.

6.1 Nihilism

The first option denies that probability applies to epistemic discourse. This option is familiar from the philosophical logic literature on conditionals and probability. Theorists like Adams 1975 and Edgington 1995 deny that conditionals have probability, though both of them preserve a weaker kind of connection between conditionals and probability. (On Adams' view, conditionals have degrees of assertability, which behave in some ways like probability. On Edgington's view, conditionals are used to express speaker's credences, though they are not the bearers of credences themselves.) Informational theorists may claim that, in a similar way, all modal claims fail to have probability.

Nihilism might seem intuitively plausible for *might*- and *must*-claims. When asked what probability they assign to *It might be raining*, a number of speakers find the question difficult to answer. But it is *prima facie* difficult to believe for conditionals. Speakers have crisp judgments that conditionals have non-trivial probabilities, and in a large number of cases these judgments line up with the corresponding conditional probabilities (see Evans and Over 2004 for an overview of relevant experimental literature). It's unclear whether nihilism has the resources to explain these facts.

To be sure, there is a familiar strategy that nihilists can appeal to: they can give a sophisticated error theory about judgments of probability of conditionals. The strategy consists in vindicating the truth of object language sentences of the form *The probability that if A, then B is n* (where *n* is the conditional probability of B, given A), while allowing that the probability of the unembedded sentence *if A, then B* may actually diverge from *n*. To pursue this strategy, we start by assuming an appropriate semantics for object language *probably*²². Following a classical idea tracing back to Kratzer²³, we let *if*-clauses work as the restrictor on the relevant information state picked out by the operator. As a result, a sentence of the form $\ulcorner A \triangleright B \urcorner$ is predicted to make a claim about the conditional probability of B, given A. For example, (10) is true just in case the probability of the coin landing tails, conditional on Frida tossing it, is .5.

(10) It is 50% likely that, if Frida tossed the coin, the coin landed tails.

²²See Yalcin 2010, 2012, Lassiter 2011, Holliday and Icard 2013 for discussion

²³To my knowledge, Kratzer never says this fully explicitly in a published work. But this point is widely attributed to her, and I have personally seen her make it in conversation.

Even assuming the restrictor maneuver, nihilists remain open to a number of challenges. First, capturing the functioning of object language probability operators might not be sufficient to capture all the data (see e.g. Khoo and Santorio 2018 for discussion of probability of conjunctions of conditionals or data involving propositional anaphora).²⁴ Second, the decision to not assign probability to conditionals is in tension with centering principles, one of which was already discussed above.

Centering. $A > B \models A \supset B$
Strong Centering. $A \wedge B \models A > B$

According to the centering principles, conditionals are entailed by and entail factual claims.²⁵ Hence the probability of a conditional is naturally bounded from above and below by two claims that unequivocally have probability. So it would be surprising if probabilities just didn't apply to conditionals.

6.2 Nonclassical probability

The second option is to hold on to informational consequence and to **Certainty Preservation**, while blocking the proof of **Informational Triviality** by giving up some other principle involved in the proof. I have already discussed **Plenitude**. I also assume that **Identity** is too basic to be questioned credibly. So I focus on the other two principles involved:

Bound. If $Pr(B | A) = 1$, then $Pr(B) \geq Pr(A)$

²⁴Here is an example about conjunction. Consider the following scenario:

Coins. Martina is considering tossing two fair coins, A and B, in two independent tosses. You leave the room before you discover whether she tosses them or not.

Now, assess the probability of the following statements:

- (i) Coin A landed heads, if it was tossed, and coin B landed tails, if it was tossed.
- (ii) Each of coin A and coin B landed heads, if it was tossed.

One natural judgment, which has been confirmed by several speakers, is that both (i) and (ii) have probability 1/4. But judgments about corresponding sentences involving overt probability operators cannot be captured via the restrictor analysis:

- (iii) It's 1/4 likely that coin A landed heads, if it was tossed, and coin B landed tails, if it was tossed.
- (iv) It's 1/4 likely that each of coin A and coin B landed heads, if it was tossed.

The reason is that, in both cases, there is an extra item (conjunction in one case, the quantifier in the other) between the probability operator and the restrictors. So the latter cannot restrict the former.

²⁵Centering for complex conditionals is controversial, as I have pointed out above, but it unequivocally holds for simple conditionals, i.e. conditionals not involving modality in the antecedent and the consequent.

Closure. For all C : If $Pr(\bullet)$ models a rational credence function, then $Pr(\bullet | C)$ also models a rational credence function.

Both principles have to do with the notion of conditional probability, which figures prominently in a theory of update. On reflection, intervening on this point seems a natural option for the informational theorist. On informational semantics, starting from Veltman 1996, *might*-claims violate a condition of Persistence, understood as follows.

Persistence. For all i : if $i \models A$, then for all $i' \subseteq i$, $i' \models A$

Persistence says that, if a formula is supported by an information state, it is supported also by any information state more informed than it. *might*-claims, as characterized by informational semantics, are an obvious counterexample to Persistence. Now, Persistence is a qualitative analog of the requirement that update preserves credence 1, which is entailed by classical conditionalization. So failures of Persistence should lead us to expect failures of update by conditionalization.²⁶

7 Conclusion

How is informational consequence linked to credence? I have argued that, plausibly, informational consequence should be regarded as certainty preserving. I.e., on any rational credence distribution, when the premises of an informational inferences have credence 1, the conclusion also has credence 1.

Unfortunately, this simple and plausible constraint generates a new kind of triviality result. The result differs from standard triviality results in the literature, in that it doesn't connect to the semantics of any particular expression. Rather, it targets directly the existence of a notion of consequence that is extensionally different from classical consequence and that preserves credence 1. It can be proved that the existence of this notion of consequence is incompatible with some basic assumptions about credence and update.

One may want to use the foregoing as an objection to informational consequence. As I have made clear, this is not my goal here. Informational consequence plays an important explanatory role and should not be given up. At the same time, informational theorists should feel pressure to give an account of credence and update that avoids triviality. The one-line moral of this paper is: a nonclassical notion of consequence needs a nonclassical account of credence and credal update.²⁷

²⁶This connection is also noticed and discussed explicitly by Bradley 2007.

²⁷*Acknowledgments suppressed for blind review.*

References

- Ernest Adams. *The logic of conditionals: An application of probability to deductive logic*, volume 86. Springer Science & Business Media, 1975.
- Author/s. Probabilities of conditionals in informational semantics. unpublished draft, 2019a.
- Author/s. Indeterminacy and triviality. unpublished draft, 2019b.
- Justin Bledin. Modus ponens defended. *Journal of Philosophy*, 112(2):57–83, 2015.
- Justin Bledin and Tamar Lando. Closure and epistemic modals. *Philosophy and Phenomenological Research*, 97(1):3–22, 2018.
- Richard Bradley. A preservation condition for conditionals. *Analysis*, 60(267): 219–222, 2000.
- Richard Bradley. A defence of the ramsey test. *Mind*, 116(461):1–21, 2007.
- Fabrizio Cariani. On Stalnaker’s *Indicative Conditionals*. forthcoming in *Studies in Linguistics and Philosophy*, Volume 100, L. McNally, Y. Sharvit and Z. Szabó (eds.), 2020.
- David Christensen. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford University Press, 2004.
- Dorothy Edgington. On conditionals. *Mind*, pages 235–329, 1995.
- Andy Egan, John Hawthorne, and Brian Weatherson. Epistemic modals in context. In G. Preyer and G. Peter, editors, *Contextualism in Philosophy*, pages 131–170. Oxford University Press, 2005.
- Jonathan St. B. T. Evans and David E. Over. *If*. Oxford University Press, Oxford, 2004.
- Hartry Field. What is logical validity? In Colin R. Caret and Ole T. Hjortland, editors, *Foundations of Logical Consequence*. Oxford University Press, 2015.
- Allan Gibbard. Two recent theories of conditionals. In William Harper, Robert C. Stalnaker, and Glenn Pearce, editors, *Ifs*, pages 211–247. Reidel, 1981.
- Anthony S. Gillies. Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616, 2004.
- Anthony S. Gillies. On truth-conditions for if (but not quite only if). *Philosophical Review*, 118(3):325–349, 2009.

- Simon Goldstein. Triviality results for probabilistic modals. *Philosophy and Phenomenological Research*, forthcoming.
- Alan Hájek and N. Hall. The Hypothesis of the Conditional Construal of Conditional Probability. In Ellery Eells, Brian Skyrms, and Ernest W. Adams, editors, *Probability and Conditionals: Belief Revision and Rational Decision*, page 75. Cambridge University Press, 1994.
- Gilbert Harman. Logic and reasoning. *Synthese*, 60(1):107–127, 1984.
- Wesley Holliday and Thomas Icard. Measure semantics and qualitative semantics for epistemic modals. In *Proceedings of SALT 23*, pages 514–534. 2013.
- David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes From Kaplan*, pages 481–563. Oxford University Press, 1989a.
- David Kaplan. Demonstratives: Afterthoughts. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes From Kaplan*, pages 565–614. Oxford University Press, 1989b.
- Justin Khoo. A note on gibbard’s proof. *Philosophical Studies*, 166(1):153–164, 2013.
- Justin Khoo and Paolo Santorio. Lecture notes: Probability of conditionals in modal semantics. Lecture notes for a course at the North American Summer School in Logic, Language and Information, Carnegie Mellon University, Summer 2018, 2018.
- Niko Kolodny. How does coherence matter? *Proceedings of the Aristotelian Society*, 107(1pt3):229–263, 2007.
- Niko Kolodny. Why be disposed to be coherent? *Ethics*, 118(3):437–463, 2008.
- Niko Kolodny and John MacFarlane. Ifs and oughts. *Journal of Philosophy*, 107(3):115–143, 2010.
- Angelika Kratzer. Conditionals. In *Chicago Linguistics Society: Papers from the Parasession on Pragmatics and Grammatical Theory*, volume 22, pages 1–15, University of Chicago, Chicago IL, 1986. Chicago Linguistic Society.
- Angelika Kratzer. *Modals and Conditionals: New and Revised Perspectives*, volume 36. Oxford University Press, 2012.
- Daniel Lassiter. *Measurement and Modality: The scalar basis of modal semantics*. PhD thesis, NYU, 2011.
- David Lewis. Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85(3):297–315, 1976.

- John MacFarlane. In what sense (if any) is logic normative for thought? unpublished draft, UC Berkeley, 2004.
- John MacFarlane. Epistemic modals are assessment-sensitive. In Andy Egan and B. Weatherson, editors, *Epistemic Modality*. Oxford University Press, 2011.
- John MacFarlane. *Assessment Sensitivity: Relative Truth and its Applications*. OUP Oxford, 2014.
- Vann McGee. A counterexample to modus ponens. *Journal of Philosophy*, 82(9): 462–471, 1985.
- Jeffrey Sanford Russell and John Hawthorne. General dynamic triviality theorems. *Philosophical Review*, 125(3):307–339, 2016.
- Moritz Schulz. Epistemic modals and informational consequence. *Synthese*, 174(3):385–395, 2010.
- Robert Stalnaker. Indicative conditionals. *Philosophia*, 5, 1975.
- Robert Stalnaker. *Inquiry*. Cambridge University Press, 1984.
- Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25(5-6):701–21, December 2002.
- Una Stojnic. One’s modus ponens: Modality, coherence and logic. *Philosophy and Phenomenological Research*, 95(1):167–214, 2017.
- Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.
- Malte Willer. A remark on iffy oughts. *Journal of Philosophy*, 109(7):449–461, 2012.
- J. Robert G. Williams. Probability and nonclassical logics. In Alan Hájek and Chris Hitchcock, editors, *Oxford Handbook to Probability and Philosophy*. Oxford University Press, 2016.
- Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, 2007.
- Seth Yalcin. Probability operators. *Philosophy Compass*, 5(11):916–937, 2010.
- Seth Yalcin. Bayesian expressivism. *Proceedings of the Aristotelian Society*, 112(2pt2):123–160, 2012.