Modus Ponens Defended

Justin Bledin*

July 5, 2016

Abstract: Is *modus ponens* valid for the indicative conditional? McGee [1985] famously presents several alleged counterexamples to this inference rule. More recently, Kolodny and MacFarlane [2010] and Willer [2010] argue that *modus ponens* is unreliable in certain hypothetical contexts. However, none of these attacks undermines an informational conception of logic on which *modus ponens* is valid.

1 Introduction

Even in logic, is nothing sacred? Intuitionistic logicians reject the law of excluded middle.[1] Quantum logicians reject the distributive law.[2] What about *modus ponens* for the indicative conditional, the hallmark of good reasoning that lets one pass from an indicative ⌜If φ then ψ⌝ and its antecedent φ to its consequent ψ? This rule has also been rejected. McGee [1985] famously argues that *modus ponens* arguments involving right-nested conditionals are invalid. Carrying the torch, Kolodny and MacFarlane [2010] argue that *modus ponens* can lead reasoners astray in hypothetical contexts triggered by supposition, and Willer [2010] also

*Published in *Journal of Philosophy* CXII(2): 57–83, 2015. For helpful comments and discussion, I am grateful to Peter Achinstein, Steven Gross, Niko Kolodny, John MacFarlane, Rob Rynasiewicz, Malte Willer, Seth Yalcin, and audiences at SEP 2014 and DEON 2014.

[1]Excluded middle holds that $\ulcorner \varphi \vee \neg\varphi \urcorner$ is valid for each sentence φ in a language $\mathcal{L}$. See Brouwer [1921] and [1923] for criticism of this law.

[2]The distributive law holds that $\ulcorner \varphi \wedge (\psi \vee \chi) \urcorner$ is equivalent to $\ulcorner (\varphi \wedge \psi) \vee (\varphi \wedge \chi) \urcorner$, and $\ulcorner \varphi \vee (\psi \wedge \chi) \urcorner$ is equivalent to $\ulcorner (\varphi \vee \psi) \wedge (\varphi \vee \chi) \urcorner$. Birkhoff and Von Neumann [1936] and Putnam [1968] are classic references on quantum logic.

argues that this rule is unreliable in certain hypothetical contexts on the basis of Thomason conditionals and Moore's Paradox.[3]

My primary aim in this paper is to defend *modus ponens* for the indicative against these three lines of attacks. I argue that none of them undermines the sanctity of this inference rule. My secondary aim is to give currency to the informational view of logic and deductive inquiry motivated in Yalcin [2007] and Bledin [2014]. As I hope to show in what follows, this informational framework affords a sharp analysis of deductive argumentation involving indicative conditionals, informational modals, and belief reports.[4]

2 First Attack

Let me begin with McGee's [1985] original attack on *modus ponens*, which involves several alleged counterexamples to this inference rule. Since the differences between these counterexamples are unimportant for present purposes, I will discuss just the most famous example (p. 462):

(P1) If a Republican wins the election, then if it is not Reagan who wins it will be Anderson.

(P2) A Republican will win the election.

(C) If it is not Reagan who wins, it will be Anderson.

The deliberative context is the period immediately preceding the 1980 U.S. presidential election, when opinion polls projected that Republican Ronald Reagan would win, with Democrat Jimmy Carter coming second and Republican John Anderson a distant third. Given these projections, McGee writes:

[3] Thomason's original example appears in van Fraassen [1980].

[4] Though I focus exclusively on the natural language indicative conditional in this paper, it is worth mentioning that *modus ponens* rules for other kinds of conditional also stand in need of defense. For instance, McGee's [1985] attack extends to *modus ponens* for *subjunctive* conditionals, and Briggs [2012] provides a causal modeling semantics that invalidates *modus ponens* for counterfactuals. Even *modus ponens* for the *material* conditional '$\supset$' has come under fire from some corners. While many logicians and philosophers think that '$\supset$' constitutively satisfies *modus ponens*—such that rejecting *modus ponens* for the indicative is tantamount to rejecting the material analysis of the indicative—Anderson and Belnap [1975], the fathers of relevance logic, reject *disjunctive syllogism* (*modus ponens* for the material conditional modulo *double negation* rules) and introduce a new relevant conditional that satisfies *modus ponens*. Dialetheists like Priest [1979] and Beall [2009] also claim that the inference from $\ulcorner\varphi \supset \psi\urcorner$ and φ to ψ can fail when both φ and $\ulcorner\neg\varphi\urcorner$ are true.

> There are occasions on which one has good grounds for believing the premises of an application of modus ponens but yet one is not justified in accepting the conclusion. (p. 462)
>
> Modus ponens is not an entirely reliable rule of inference. 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. (p. 463)

Since Reagan and Anderson are the only Republican candidates, one is clearly justified in, or has good grounds for, believing that (P1) is true.[5] The polls strongly suggest that Reagan will win the election, so one is also presumably justified in believing that (P2) is true. However, the polls strongly suggest that Carter will get more votes than Anderson, so McGee claims that one is unjustified in believing that (C) is true.

On Stalnaker's [1968] and [1975] seminal semantics for the indicative conditional, *modus ponens* is valid since this rule unrestrictedly preserves truth.[6] So McGee's example also purportedly shows that Stalnaker's semantics is incorrect. Indeed, McGee presents a modified version of the Stalnakerian semantics that invalidates *modus ponens*; the full weight of his attack does not rest solely on a few examples.[7]

To see this, consider a formal language $\mathcal{L}$ with the following symbols: sentential atoms like 'A', 'B', 'C', ..., the contradiction symbol '$\bot$', the sentential connectives '$\neg$', '$\wedge$', and '$\vee$', the indicative conditional '$\Rightarrow$', and parentheses. Assume that $\mathcal{L}$ has the usual grammar. Let $At_{\mathcal{L}}$ and $S_{\mathcal{L}}$ designate the set of atoms and well-formed sentences of $\mathcal{L}$ respectively.

Our starting point is the following Stalnakerian semantics:[8]

---

[5]Throughout this essay, I use the truth predicate as merely a disquotational and quantificational device. I do not look to challenge philosophers like Adams [1975] and Edgington [1986] who think that indicative conditional declarative sentences lack truth conditions.

[6]In the earlier [1968] paper, subjunctive conditionals play the starring role, but Stalnaker aims to provide a unified theory of indicatives and subjunctives.

[7]Unfortunately, McGee's semantics is all but ignored in critical discussion of his paper. Katz [1999] even expresses skepticism about the possibility of such a semantics:

> I find it doubtful that there is any consistent interpretation of the conditionals occurring in McGee's example, truth functional or otherwise, which will make the premises true and the conclusion false. (p. 405-6)

[8]My presentation differs from Stalnaker's [1968] original presentation in some minor respects: his models include a Kripkean accessibility relation $\mathcal{R}$ between worlds whereas mine do not, my models include the selection function f whereas he takes this function to be part of the semantic apparatus outside the model structure, his

Def 1. A *model* $\mathcal{M} = \langle \mathcal{W}, \mathcal{V}, f \rangle$ for $\mathcal{L}$ consists of a nonempty set of worlds $\mathcal{W}$, an interpretation function $\mathcal{V}$ that maps each sentential atom $p \in At_{\mathcal{L}}$ and world $w \in \mathcal{W}$ to a truth value, and a *selection function* f that maps each sentence $\varphi \in \mathcal{S}_{\mathcal{L}}$ and *base world* $w \in \mathcal{W}$ to a *selected world* $v \in \mathcal{W}$.

Since $f(\varphi, w)$ is intended to return the φ-world that differs minimally from w according to some similarity metric, the selection function must meet certain minimal constraints.[9] For our purposes, the most important constraint is this: if φ holds in w, then $f(\varphi, w) = w$. That is, the most similar φ-world to a base world in which φ holds is the base world itself.

Def 2. *Truth at a world* in model $\mathcal{M}$ is recursively defined for the full language $\mathcal{L}$ as follows:[10]

$$
\begin{array}{rll}
p \text{ is true at } w & \text{iff} & \mathcal{V}(p, w) = T \\
\text{`}\bot\text{' is true at } w & \text{iff} & 0 = 1 \\
\ulcorner \neg\varphi \urcorner \text{ is true at } w & \text{iff} & \varphi \text{ is false at } w \\
\ulcorner \varphi \wedge \psi \urcorner \text{ is true at } w & \text{iff} & \varphi \text{ and } \psi \text{ are true at } w \\
\ulcorner \varphi \vee \psi \urcorner \text{ is true at } w & \text{iff} & \varphi \text{ or } \psi \text{ is true at } w \\
\ulcorner \varphi \Rightarrow \psi \urcorner \text{ is true at } w & \text{iff} & \psi \text{ is true at } f(\varphi, w)
\end{array}
$$

The selection function f comes into play in the final semantic clause for the indicative. For example, 'If Reagan wins then Republicans will rejoice' is true at w just in case Republicans rejoice in the closest world to w where Reagan wins according to the notion of similarity encoded in f.

Def 3. The argument from $\varphi_1, ..., \varphi_n$ to ψ is *valid*, $\{\varphi_1, ..., \varphi_n\} \models_0 \psi$, just in case there is no world $w \in \mathcal{W}$ such that each of $\varphi_1, ..., \varphi_n$ is true at w but ψ is false at w.

Given the Stalnakerian recursive semantics and this formal consequence relation, it is easy to see that *modus ponens* for the indicative is valid. Suppose that both $\ulcorner \varphi \Rightarrow \psi \urcorner$ and φ are true at w. Then ψ is true at $f(\varphi, w)$ and $f(\varphi, w) = w$, so ψ is true at w. Thus, $\{\ulcorner \varphi \Rightarrow \psi \urcorner, \varphi\} \models_0 \psi$.

---

selection function sends *propositions* in $2^{\mathcal{W}}$ and base worlds in $\mathcal{W}$ to selected worlds in $\mathcal{W}$, and so forth. However, these differences are unimportant for present purposes.

[9] The similarity metric is supplied by context. We could make this detail explicit by adding a set of contexts $\mathcal{C}$ to our models where each $c \in \mathcal{C}$ determines a selection function f_c. However, I want to keep the formal semantics in this paper relatively simple.

[10] For a cleaner exposition, I will often leave implicit the relativization of truth at a world to a model $\mathcal{M}$. I also ignore conditionals with impossible antecedents. To handle these, Stalnaker's models also include an absurd world λ where every $\varphi \in S_{\mathcal{L}}$ is true at λ. Indicative conditionals with impossible antecedents are vacuously true.

But if McGee's argument from (P1) and (P2) to (C) is a genuine counterexample to *modus ponens*, where exactly does Stalnaker go wrong? Let '*Rep*', '*R*', and '*A*' abbreviate 'A Republican wins', 'Reagan wins', and 'Anderson wins' respectively, and assume that $\mathcal{W}$ includes worlds w_R, w_C, and w_A in which Reagan, Carter, or Anderson wins the election respectively, where w_R is actual and w_C is much more similar to w_R than w_A is, in the context where the opinion polls are being considered. Then, as McGee claims, (P2) is true and (C) is false: '*Rep*' is true at w_R and since $f(`\neg R', w_R) = w_C$, '$\neg R \Rightarrow A$' is false at $w_R$. However, (P1) is false: since $f(`Rep', w_R) = w_R$, '$Rep \Rightarrow (\neg R \Rightarrow A)$' and (C) have the same truth value at w_R.

The problem with the Stalnakerian semantics, it seems, is that if φ is true at w, then $\ulcorner \varphi \Rightarrow (\psi \Rightarrow \xi) \urcorner$ and $\ulcorner \psi \Rightarrow \xi \urcorner$ rise and fall together. But McGee wants the extensions of (P1) and (C) to come apart. To break their equivalence, he suggests that we modify the semantics slightly and evaluate each sentence $\varphi \in S_{\mathcal{L}}$ for truth at a world $w \in \mathcal{W}$ *under a set of hypotheses* $\Gamma \subseteq S_{\mathcal{L}}$. More specifically, he proposes the following revised compositional semantics for $\mathcal{L}$, where f now selects the most similar world to w at which the sentences in hypothesis set Γ all hold.[11]

Def 4. *Truth at a world under* Γ is recursively defined as follows:

$$
\begin{array}{rcl}
p \text{ is true at } w \text{ under } \Gamma & \text{iff} & \mathcal{V}(p, f(\Gamma, w)) = T \\
`\bot' \text{ is true at } w \text{ under } \Gamma & \text{iff} & 0 = 1 \\
\ulcorner \neg\varphi \urcorner \text{ is true at } w \text{ under } \Gamma & \text{iff} & \varphi \text{ is false at } w \text{ under } \Gamma \\
\ulcorner \varphi \wedge \psi \urcorner \text{ is true at } w \text{ under } \Gamma & \text{iff} & \varphi \text{ and } \psi \text{ are true at } w \text{ under } \Gamma \\
\ulcorner \varphi \vee \psi \urcorner \text{ is true at } w \text{ under } \Gamma & \text{iff} & \varphi \text{ or } \psi \text{ is true at } w \text{ under } \Gamma \\
\ulcorner \varphi \Rightarrow \psi \urcorner \text{ is true at } w \text{ under } \Gamma & \text{iff} & \psi \text{ is true at } w \text{ under } \Gamma \cup \{\varphi\}
\end{array}
$$

An indicative conditional $\ulcorner \varphi \Rightarrow \psi \urcorner$ is true at w under the hypothesis set Γ just in case its consequent ψ is true at w under the augmented hypothesis set $\Gamma \cup \{\varphi\}$ obtained by adding the antecedent φ to Γ. Thus, if φ is true at world w *simpliciter* just in case φ is true at w under the empty hypothesis set $\emptyset$, then the argument from (P1) and (P2) to (C) fails to preserve truth at w_R. Since $f(\{`Rep', `\neg R'\}, w_R) = w_A$, '*A*' is true at w_R under $\{`Rep', `\neg R'\}$, so '$Rep \Rightarrow (\neg R \Rightarrow A)$' is true at w_R under $\emptyset$. Also, '*Rep*' is true at w_R under $\emptyset$. However, since $f(\{`\neg R'\}, w_R) = w_C$, '$\neg R \Rightarrow A$' is false at w_R under $\emptyset$. *Modus ponens* for the indicative fails.

[11] The selection function must meet at least the following constraint: $f(\Gamma, w) = w$ if each member of Γ holds in w. For ease of exposition, I ignore the possibility of inconsistent hypothesis sets.

3 First Defense

Time for my first defense. At this point, it will be helpful to distinguish between the following kinds of argument:

(I) deductively good argument

(II) logically valid argument

(III) argument that is normative for belief

As I use the term, 'deductively good argument' is an evaluative concept applicable inside the third-person standpoint of appraisal. Deductively good arguments are those that we can competently make, by virtue of logical form, in both categorical deliberative contexts and hypothetical contexts triggered by supposition—at least, we can competently make these arguments in any deliberative context where doing so is not simply a waste or misuse of scarce cognitive resources.[12] Publicly, if someone asserts or supposes that each premise of a deductively good argument is true, then we can infer on this basis that its conclusion is true. Privately, if you activate your beliefs or simply suppose in an episode of internal theoretical deliberation that the premises are true, then you can infer that the conclusion is true.[13]

Once logicians and philosophers start talking about 'logical validity', they are typically theorizing about some objective feature of arguments that makes them deductively good.[14] On the standard truth preservation view, an argument is logically valid if and only if it is impossible for each of its premises to be true and for its conclusion to be false by virtue of their logical form.[15] This informal characterization of validity leaves room for much disagreement (What sense of 'impossible' is relevant?

[12]As Peter Achinstein points out (p.c.), the argument from 'A Republican wins' to 'A Republican wins' is good but we rarely, if ever, do well to make it.

[13]I am using 'infer' in a thin sense. Inference consists in recognizing what follows; it need not culminate in belief formation.

[14]Tarski [1936a] opens his seminal essay on logical consequence with this remark on the subject of mathematical logic:

> The concept of logical consequence is one whose introduction into the field of strict formal investigation was not a matter of arbitrary decision on the part of this or that investigator; in defining this concept, efforts were made to adhere to the common usage of the language of everyday life. (p. 409)

I am doubtful that such a "common usage" of logical consequence exists. Though people on the street have intuitions about what follows from what, and about what we can competently infer in our deliberations, the idea that 'logically valid' arguments necessarily preserve truth by virtue of logical form is a theoretical idea that entered philosophy and mathematics with Aristotle.

[15]The formal relation $\models_0$ in Def 3 is meant to explicate this informal concept.

What counts as 'logical form'?) and is itself open to dispute. Indeed, I will soon introduce an alternative *informational* view of validity. For now, let me just note that while logical validity and deductively good argument presumably coincide—an argument is valid just in case we can competently make it in our deliberations by virtue of logical form—these two notions are conceptually distinct. Logical validity, as I am understanding it here, says nothing about what we can and cannot infer inside categorical and hypothetical contexts.[16]

It is commonly thought that logically valid arguments play some kind of special normative role in our epistemic practices. From the first-person standpoint of deliberation and the related second-person standpoint of advice—which is an attempt to aid in another's deliberation—logically valid arguments inform, in some special sense, what we ought and ought not to believe. However, the normative import of logic for belief is far from clear. As Harman stresses in *Change in View* [1986], logic is not a theory of reasoning, at least if reasoning is broadly conceived as a non-monotonic organic process of forming, maintaining, and revising beliefs over time—what he calls "reasoned change in view"—and not as just the churning out of consequences from a set of premises—what I have been calling 'argument' or 'inference'. Bridge principles connecting logical relations with requirements and permissions of epistemic reason are needed at the logic-epistemology interface.

Harman himself argues that various candidate bridge principles are all problematic, so he negatively concludes that logic plays no special normative role in our deliberations. However, many philosophers—for example, Broome [1999], MacFarlane [*ms.*], and Field [2009a]—have thought otherwise and endorsed bridge principles connecting logic with reasoned change in view.[17] In fact, I will shortly argue that McGee's attack on *modus ponens* for the indicative crucially hangs on a kind of logical-evidential closure principle. But before that, let me continue with a few more preliminaries. Since my agenda in this section is simply to defend *modus ponens* against McGee, I next want to sketch my preferred

[16] Some will disagree. Hartry Field [2006], [2008], [2009a], [2009b], [*ms.*], for one, argues that logical validity is a primitive notion that directly governs our inferential practices. He does not seem to think that validity and good deductive argument are entirely distinct. But let me put aside Field's radical conception of logic here. See Bledin [2014] for some discussion.

[17] To give a sense of the candidate principles, here is MacFarlane's [*ms.*] preferred closure requirement on full belief:

Closure: If schema S is valid and you apprehend the inference from $\varphi_1, ..., \varphi_n$ to ψ as an instance of S, then you ought to see to it that if you believe that each of $\varphi_1, ..., \varphi_n$ is true, then you believe that ψ is true.

conception of logic on which this inference rule comes out valid. I then argue that McGee's attack, as it stands, does not undermine this way of seeing things.

Inspired by Yalcin [2007], I develop an *informational* conception of logic and deductive inquiry in Bledin [2014]. Logic, on this view, is a science that is fundamentally concerned not with the preservation of truth but rather with the preservation of structural properties of the bodies of information that we generate, encounter, absorb, and exchange as we interact with one another and learn about our world.

At a high level of abstraction, we can think of information as that which rules out various ways the world might be while leaving others open. The information in a weather report might rule out the possibility that it will rain in Sydney but leave open the possibility that it will snow in Toronto. The informational content of many of our belief states rules out the possibility that Napoleon was taller than the average Frenchman of his time.[18] And so forth.

A body of information *incorporates* that such and such, or is a body of information according to which such and such, just in case it has a particular structure. The weather report incorporates that it will *not* rain in Sydney since it rules out the possibility that it will rain in Sydney. The weather report incorporates that it will be both sunny in Dubai *and* overcast in London just in case it rules out both the possibility that it will not be sunny in Dubai and also the possibility that it will not be overcast in London. The report is information according to which it will be either sunny in Dubai *or* overcast in London just in case it rules out the possibility that it will be neither sunny in Dubai nor overcast in London. And so forth.

Turning to conditionals, the weather report incorporates, say, that if the hurricane hits Florida then there will be widespread destruction just in case the body of information that minimally deviates from this report so as to incorporate that the hurricane will hit Florida also incorporates that there will be widespread destruction. This *minimal change state* is simply the body of information obtained by tentatively adding the information that the hurricane will hit Florida to the weather report—that is, this updated information rules out the same possible states of the world as the report but also rules out the possibility that the hurricane will bypass Florida. The report incorporates the indicative conditional just in case any state of the world still open after the update is one in

[18] Apparently, Napoleon Bonaparte was taller than average. I use 'information' in a broad enough sense to include what is sometimes labeled '*mis*information'. Non-factual information can rule out the actual state of the world.

which there will be widespread destruction.

More precisely now, the informational concept of validity preserves not truth at worlds but incorporation by bodies of information. An argument is *logically valid* if and only if any body of information that incorporates all of its premises also incorporates its conclusion by virtue of logical form. If we think of deductive inquiry as an information-driven enterprise in which a reasoner investigates what is so according to a body of information that incorporates the premises of an argument, then it is hardly surprising that validity and deductively good argument coincide. The reliable inferences that the agent can draw are precisely those that preserve incorporation.

Understanding logic along these informational lines, *modus ponens* for the indicative conditional comes out logically valid. If information incorporates an indicative, then the minimal change state incorporating its antecedent also incorporates its consequent. However, if the original information also incorporates the antecedent, then the minimal change state is just the original information itself. Thus, the original information incorporates the consequent.

What about McGee's elegant formal semantics that invalidates *modus ponens*?[19] Well, on the informational view of logic, the formal relation $\models_0$ in Def 3 that preserves truth at a world is problematic. McGee, along with Stalnaker and many others for that matter, misidentifies the informal target notion of validity. We are after a consequence relation that explicates not an informal notion of unrestricted truth preservation but rather an informal notion of unrestricted incorporation preservation.

To formalize the informational picture, we can follow Yalcin [2007] and Kolodny and MacFarlane [2010] and evaluate sentences in $S_{\mathcal{L}}$ for truth relative to a *point of evaluation* $\langle w, i\rangle$ consisting both of a world $w \in \mathcal{W}$ and of an *information state* $i \in 2^{\mathcal{W}}$.

Def 5. *Truth at a point* is recursively defined as follows:[20]

[19] I should mention that McGee does not wholeheartedly endorse this semantics. He floats the idea that an entirely new approach to conditionals might be needed, and even suggests that the logic of conditionals might not be amenable to systematic treatment.

[20] Note that the selection function f no longer has any work to do. A simpler model $\mathcal{M} = \langle \mathcal{W}, \mathcal{V}\rangle$ consisting only of a nonempty set of worlds $\mathcal{W}$ and of an interpretation function $\mathcal{V}$ is all we need.

$$\begin{array}{rcl}
p \text{ is true at } \langle w,i\rangle & \text{iff} & \mathcal{V}(p,w)=T \\
`\bot\text{' is true at } \langle w,i\rangle & \text{iff} & 0=1 \\
\ulcorner\neg\varphi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ is false at } \langle w,i\rangle \\
\ulcorner\varphi\wedge\psi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ and } \psi \text{ are true at } \langle w,i\rangle \\
\ulcorner\varphi\vee\psi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ or } \psi \text{ is true at } \langle w,i\rangle \\
\ulcorner\varphi\Rightarrow\psi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \psi \text{ is true at } \langle v,i+\varphi\rangle \text{ for all } v\in i+\varphi
\end{array}$$

where the information state $i+\varphi$ appearing in the final clause for the indicative conditional is the largest subset $i' \subseteq i$ such that φ is true at $\langle w,i'\rangle$ for all $w \in i'$.[21]

A sentential atom $p \in At_{\mathcal{L}}$ is true at point $\langle w,i\rangle$ just in case the world parameter w is a p-world. The clauses for '$\bot$', '$\neg$', '$\wedge$', and '$\vee$' are all straightforward. But the recursive clause for the indicative conditional '$\Rightarrow$' is more interesting. Note that every sentence $\varphi \in S_{\mathcal{L}}$ corresponds to a structural constraint on information states:

Def 6. Information state i *incorporates* φ, $i \rhd \varphi$, just in case φ is true at $\langle w,i\rangle$ for all $w \in i$.

Stated in terms of incorporation, $\ulcorner\varphi \Rightarrow \psi\urcorner$ is true at $\langle w,i\rangle$ just in case the largest subset of the information state parameter i that incorporates the antecedent of this indicative also incorporates its consequent: $i+\varphi \rhd \psi$.

The set $i+\varphi$ models the minimal change state mentioned above in my informal discussion of conditionals. The technical binary relation $\rhd$ in Def 6 holds between a sentence $\varphi \in S_{\mathcal{L}}$ and an information state $i \in 2^{\mathcal{W}}$ if and only if information modeled by i is information according to which φ is true. Using this formal notion of incorporation, we can replace the formal consequence relation in Def 3 with the following explication of the informational concept of validity:

Def 7. The argument from $\varphi_1, ..., \varphi_n$ to ψ is *valid*, $\{\varphi_1, ..., \varphi_n\} \models_I \psi$, just in case there is no information state $i \in 2^{\mathcal{W}}$ such that $i \rhd \varphi_1$,...,$i \rhd \varphi_n$ but $i \not\rhd \psi$.

This new consequence relation—what Yalcin [2007] calls "informational consequence"[22]—validates *modus ponens* for the indicative conditional.

[21] First aside: The idea that 'if'-clauses restrict quantificational operators goes back to Lewis [1975]. Second aside: Yalcin [2007] requires that $i+\varphi \neq \emptyset$ but I follow Kolodny and MacFarlane [2010] and relax this restriction. Third aside: When dealing with indicatives more complex than those considered in this paper, the semantic clause provided is too simple since there need not be a unique *maximal φ-subset* $i' \subseteq i$ such that both φ is true at $\langle w,i'\rangle$ for all $w \in i'$ and for each i'' such that $i' \subset i'' \subseteq i$, φ is false at $\langle w,i''\rangle$ for some $w \in i''$. See Kolodny and MacFarlane [2010] for a more sophisticated clause for the general case.

[22] Veltman [1996] proposes a similar consequence relation defined over his dynamic update semantics. Kolodny and MacFarlane [2010] also consider a variant of $\models_I$ but

Suppose that $i \rhd \ulcorner\varphi \Rightarrow \psi\urcorner$ and $i \rhd \varphi$. Then $i + \varphi \rhd \psi$ and $i = i + \varphi$, so $i \rhd \psi$. Thus, $\{\ulcorner\varphi \Rightarrow \psi\urcorner, \varphi\} \models_I \psi$.

So, what about McGee's apparent counterexample to *modus ponens*? Doesn't this show that the informational view cannot be correct? An informational logician might, of course, object to the epistemological assumptions that give this 'counterexample' purchase. For instance, one might deny that an agent is unjustified in believing the conclusion of the argument on the basis of the opinion polls.[23] However, let us grant that an agent apprised of the polls has good grounds for believing that if a Republican wins then if it is not Reagan who wins it will be Anderson, and for believing that a Republican will win the election, but this agent is unjustified in believing that if it is not Reagan who wins it will be Anderson. McGee's election example still does not undermine *modus ponens*—at least not without further elaboration.

The distinction between deductively good argument, logically valid argument, and deductive argument with normative import is helpful here. I have claimed that *modus ponens* for the indicative is logically valid since it unrestrictedly preserves incorporation: (P1) and (P2) imply (C) because any body of information (the content of the opinion polls, the evening news, and so on) according to which if a Republican wins then if it is not Reagan who wins it will be Anderson, and according to which a Republican will win the election, is therefore, by virtue of logical form, also information according to which if it is not Reagan who wins it will be Anderson. *Modus ponens* is also a deductively good form of argument: in deliberative contexts where (P1) and (P2) are incorporated by the information that drives one's inquiry, one can infer (C) from this

they do not endorse this relation as the one true explication of logical consequence. More on this in §4.

[23]Sinnott-Armstrong, Moor, and Fogelin [1986] and Fulda [2010] point out that an agent has good grounds for believing that (C) is true when this conditional is analyzed as the material conditional. But I do not subscribe to the material analysis of the indicative conditional.

basis.[24]

However, I can say all of this and still concede that the argument from (P1) and (P2) to (C) fails to preserve justification. It is important to recognize that McGee's attack on *modus ponens* via counterexample presupposes a kind of bridge principle linking logic and epistemology. Specifically, his argument might seem to rely on the following principle:[25]

($\star$) If an argument is logically valid, and one has good grounds for believing each of its premises, then one also has good grounds for believing its conclusion.

Since an agent can have good grounds for believing both that (P1) is true and that (P2) is true, but still be unjustified in believing that (C) is true, the argument from (P1) and (P2) to (C) is invalid by ($\star$).

Of course, the principle ($\star$) is suspect for familiar reasons. One can arguably have good grounds for believing the many claims in a history text but still be unjustified in believing that this volume is error-free (Makinson [1965]). Moreover, one can arguably have good grounds for believing that each ticket in a large lottery will lose but still be unjustified in believing that all of the tickets will lose (Kyburg [1961]).[26]

But never mind this. The following less controversial bridge principle could also do the trick:

($\star\star$) If an argument is logically valid, and one has good grounds for

---

[24] In §5 and §7, I defend my claim that *modus ponens* is a good argument form against Kolodny and MacFarlane [2010] and Willer [2010] who argue that we can go wrong in using this rule in certain hypothetical contexts. If you are not yet convinced that McGee's *modus ponens* argument is good, consider this variant with anaphoric pronouns in the conclusion (cf. Gillies [2004]):

(P1) If a Republican wins the election, then if he is not Reagan, he is Anderson.

(P2) A Republican will win the election.

(C) If he is not Reagan, he is Anderson.

Since 'he' in (C) refers back to the Republican who will win the election in (P2), this argument might strike you as better. However, the difference between this example and McGee's original example is just a matter of style.

[25] See Sinnott-Armstrong, Moor, and Fogelin [1986], and Piller [1996] for related discussion.

[26] In addition to these Preface and Lottery cases, ordinary non-paradoxical cases threaten the bridge principle ($\star$). Suppose that φ_1 and φ_2 logically imply φ_3 and consider a toy epistemology where e is a probabilistic evidential support measure, and one has good grounds for believing that φ is true just in case $e(\varphi) \geq 0.8$. If $e(\varphi_1) = e(\varphi_2) = 0.8$, then one would be justified in believing both that φ_1 is true and that φ_2 is true. However, unless $e(\ulcorner \varphi_1 \vee \varphi_2 \urcorner) = e(\varphi_1) = e(\varphi_2) = 0.8$, it is consistent with the Kolmogorov probability axioms that $e(\varphi_3) < 0.8$ and one would be unjustified in believing that φ_3 is true.

> believing the *conjunction* of its premises, then one also has good grounds for believing its conclusion.

In a deliberative context where one is considering the opinion polls, one knows with certainty that (P1) is true, so one has good grounds for believing that the conjunction of (P1) and (P2) is true. However, one is unjustified in believing that (C) is true, so the argument from (P1) and (P2) to (C) is invalid by ($\star\star$).

Laid out in this way, McGee's election example establishes only that we must either abandon *modus ponens* for the indicative conditional or reject ($\star\star$). The logical rule and evidential closure principle cannot both hold. Which should we give up? Admittedly, the bridge principle ($\star\star$) is, at first glance, highly plausible.[27] Something like this principle might be thought to underwrite the idea that logical deduction is an epistemically secure means to extend one's beliefs. As such, it is hardly surprising that McGee's example has received so much attention in the literature and led many philosophers to seriously question the validity of *modus ponens*. However, if the informational view of logic is correct, then ($\star\star$) must go. McGee's example is not a genuine counterexample to *modus ponens* but rather a counterexample to a certain natural tie between logical validity and epistemic reason.

From the informational perspective, such counterexamples to ($\star\star$) should, after some reflection, be expected. On the informational view of logic, an argument is valid just in case every body of information, including one's total evidence, that incorporates each of the argument's premises also incorporates its conclusion by virtue of logical form. But having good grounds for belief does not preclude a risk of inaccuracy. One might be justified in believing the conjunction of an argument's premises even if one's total evidence does not incorporate them. In fact, one might be justified in believing this conjunction even if one's evidence does not incorporate the argument's conclusion either. This arguably opens the door for justification loss. For some deductive arguments, it might further be the case that one's total evidence provides insufficient grounds for believing the conclusion, so principle ($\star\star$) fails.

Those who think that McGee's election argument fails to preserve justification from its conjoined premises to its conclusion will presumably think that the following valid single-premise argument also fails to

[27] Well, a suitably qualified version of ($\star\star$) is highly plausible. Presumably, a mathematician can have good grounds for believing the conjunction of the axioms of Peano Arithmetic but still have insufficient grounds for believing some complex unproven theorem of PA. We might restrict the bridge principle to sufficiently obvious or recognized logical entailments.

preserve justification:[28]

(P1) It is not the case that a Democrat will win the election.

(C) It is not the case that a Democrat might win the election.

Information incorporating (P1) rules out the possibility that a Democrat will win, and therefore it is also information incorporating that it is not the case that a Democrat might win. However, while someone apprised of the opinion polls right before the 1980 U.S. presidential election would be justified in believing that (P1) is true, this person would arguably be unjustified in believing that (C) is true.

Further, this argument is logically valid on the informational view:

(P1) Either Reagan or Carter will win the election.

(P2) Carter will not win the election.

(C) It must be the case that Reagan will win the election.

Information incorporating both (P1) and (P2) rules out the possibility that Reagan will not win, and therefore it is also information according to which it must be the case that Reagan will win. However, consider an agent who would be justified in believing that Reagan will win on the basis of evidence that does not rule out the possibility that he will not win. Presumably, this agent would also be justified in believing that the conjunction of (P1) and (P2) is true. But this agent would arguably still be unjustified in believing that (C) is true. So principle ($\star\star$) is again in trouble and *modus ponens* has nothing to do with it.[29]

---

[28] As Yalcin [2007] points out, Łukasiewicz [1930] appears to endorse this form of argument (at least for hypothetical contexts).

[29] Other potential counterexamples to ($\star\star$) arise in the debate over the closure of knowledge under known entailment. For example, consider the following argument from Dretske [2005]:

(P1) There are cookies in the jar.

(P2) If there are cookies in the jar, then there is a mind-independent physical reality.

(C) There is a mind-independent physical reality.

Dretske argues that one's perceptual reasons for believing that there are cookies in the jar do not transmit to the "heavyweight implication" (C). Thus, one can have good grounds for believing that the conjunction of (P1) and (P2) is true but have insufficient grounds for believing that (C) is true.

That said, the status of this example is highly controversial. Contextualists about epistemic justification, for instance, hold that one can justifiably believe that (C) is true in low-standard contexts, but one cannot even justifiably believe that (P1) is true in high-standard skeptical contexts.

4 Second Attack

So much for the first attack. *Modus ponens* for the indicative conditional, and the informational concept of logical validity that upholds it, is left unscathed by McGee's argument. However, my defense of *modus ponens* is not yet complete. Let me now turn to more recent arguments that purportedly show how this inference rule is unreliable in hypothetical deliberative contexts. These attacks directly challenge my claim that *modus ponens* is a deductively good form of argument.

Kolodny and MacFarlane [2010] consider a scenario where ten miners are trapped inside either shaft A or shaft B, but you do not know which.[30] Floodwaters are approaching these shafts. Armed with some sandbags, you can block either shaft A or shaft B, but not both. If you block the shaft with the miners, then all ten live. If you block the shaft without the miners and divert the water into the other shaft containing them, then all ten die. If you block neither shaft, then the water will flow into both shafts, killing only the single miner who is lowest down. What should you do?

Given your ignorance about the miners' whereabouts, Kolodny and MacFarlane state that you ought to block neither shaft. But you might reason as follows:[31]

1	The miners are in shaft A or shaft B	Premise
2	If the miners are in A, then I ought to block A	Premise
3	If the miners are in B, then I ought to block B	Premise
4	The miners are in shaft A	Supposition
5	I ought to block A	From 2,4
6	I ought to block A or I ought to block B	From 5
7	The miners are in shaft B	Supposition
8	I ought to block B	From 3,7
9	I ought to block A or I ought to block B	From 8
10	I ought to block A or I ought to block B	From 1,4-6,7-9

In the categorical context where you are considering whether to use the sandbags, you know that either the miners are in shaft A or the miners

[30] They credit this example to Derek Parfit who in turn credits Donald Regan.

[31] See Barwise and Etchemendy [1999] for a good introduction to Fitch-style proofs.

are in shaft B, and it seems reasonable to reflect both that if the miners are in shaft A then you ought to block A and that if the miners are in shaft B then you ought to block B. You then enter a hypothetical context by supposing that the miners are in shaft A. Applying *modus ponens* in this hypothetical context, you infer that you ought to block shaft A, and therefore, by *disjunctive weakening*, that either you ought to block A or you ought to block B. Similarly, you suppose that the miners are in shaft B and you infer inside the triggered hypothetical context that this same disjunction holds. Finishing the *proof by cases*, you conclude that either you ought to block A or you ought to block B.

But where does this reasoning go wrong? Kolodny and MacFarlane consider a number of ways to resist its conclusion without abandoning or restricting the classical rules of inference used in the proof—rejecting one or more of the premises, disambiguating between different senses of 'ought', or taking the two conditional premises to have a non-obvious logical form. However, they argue that none of these escape routes work. Moreover, *modus ponens* for the indicative conditional seemingly leads to trouble in other episodes of deliberation that do not involve either *disjunctive weakening* or *proof by cases* (more on this in §5). So Kolodny and MacFarlane ultimately pin the blame on *modus ponens*. In at least some hypothetical contexts, they conclude, this inference rule can lead you astray.

Kolodny and MacFarlane, like McGee, bolster their attack with a formal semantics that invalidates *modus ponens*. We can capture the relevant aspects of their semantics by adding *epistemic* necessity '$\Box_e$' and possibility '$\Diamond_e$' modals and *deontic* necessity '$\Box_d$' and possibility '$\Diamond_d$' modals to the language $\mathcal{L}$ from §2 and supplementing the compositional semantics in Def 5 with recursive clauses for these informational modals.

Def 8. A *model* $\mathcal{M} = \langle \mathcal{W}, \mathcal{V}, d \rangle$ for the expanded modal language $\mathcal{L}$ consists of a nonempty set of worlds $\mathcal{W}$, an interpretation function $\mathcal{V}$ for the sentential fragment of $\mathcal{L}$, and a *deontic selection function* d that maps each information state $i \in 2^{\mathcal{W}}$ to the set of *deontically ideal* worlds $d(i) \subseteq i$ relative to this state.

Kolodny and MacFarlane want to remain fairly neutral on the specifics of deontic ideality.[32] As we will see shortly, all that is required for their argument is that the deontically ideal worlds, if any, relative to a body of information that leaves open the possibilities both that the miners are in shaft A and that they are in shaft B are worlds in which neither

[32]Context will typically supply the function d and, more generally, will modulate the flavor of modality (epistemic, deontic, bouletic, etc.) that we are concerned with in the first place.

shaft is blocked, and the deontically ideal worlds, if any, relative to information that leaves open the possibility only that the miners are in some particular shaft, say shaft A, are ones in which this shaft is blocked.

Def 9. *Truth at a point* is defined by adding the following clauses to the semantics in Def 5:

$$\begin{array}{lll}
\ulcorner\Box_e\varphi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ is true at } \langle v,i\rangle \text{ for all } v\in i \\
\ulcorner\Diamond_e\varphi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ is true at } \langle v,i\rangle \text{ for some } v\in i \\
\ulcorner\Box_d\varphi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ is true at } \langle v,d(i)\rangle \text{ for all } v\in d(i) \\
\ulcorner\Diamond_d\varphi\urcorner \text{ is true at } \langle w,i\rangle & \text{iff} & \varphi \text{ is true at } \langle v,d(i)\rangle \text{ for some } v\in d(i)
\end{array}$$

The informational modals are effectively quantifiers over possible worlds.[33] 'The miners must be in shaft A' is true at point $\langle w,i\rangle$ just in case all of the worlds in i are worlds in which the miners are in shaft A. 'Shaft A ought to be blocked' is true at $\langle w,i\rangle$ just in case all of the deontically ideal worlds in $d(i)$ are worlds in which shaft A is blocked. And so forth.

Although Kolodny and MacFarlane consider a slight variant of the informational consequence relation $\models_I$ in Def 7, they call arguments that preserve incorporation "quasi-valid." To count as "valid," an argument must also preserve truth across all "proper points" $\langle w,i\rangle$ where $w\in i$ that correspond to an actual or possible agent's situation and knowledge state:[34]

Def 10. The argument from $\varphi_1, ..., \varphi_n$ to ψ is *valid*, $\{\varphi_1, ..., \varphi_n\} \models_{Tr} \psi$, just in case there is no proper point $\langle w,i\rangle$ such that each of $\varphi_1, ..., \varphi_n$ is true at $\langle w,i\rangle$ but ψ is false at $\langle w,i\rangle$.[35]

The guiding thought here is that whereas valid arguments are reliable in both categorical and hypothetical deliberative contexts, quasi-valid but invalid arguments are reliable in categorical contexts where the premises are known but problematic in hypothetical contexts. For instance, consider the following *modus ponens* argument from the miners example:

(P1) If the miners are in A, then I ought to block A.

(P2) The miners are in shaft A.

(C) I ought to block A.

---

[33]The basic idea that modals quantify over a contextually determined domain of possible worlds goes back to Kratzer [1977].

[34]If every context $c\in\mathcal{C}$ supplies a factual information state i_c that includes the world w_c of this context, then valid arguments must preserve truth across only those points determined by a context (cf. Kaplan [1989]).

[35]It is easy to verify that $\models_I$ and $\models_{Tr}$ coincide over the fragment of $\mathcal{L}$ without informational modals and the indicative.

This argument preserves incorporation but fails to preserve truth across all proper points. Let 'InA' and 'BlA' abbreviate 'The miners are in shaft A' and 'Shaft A is blocked', respectively. Now, consider a proper point $\langle w, i\rangle$ consisting both of a world w in which the miners are in A and of an information state i that includes worlds in which the miners are in A, worlds in which the miners are in B, and at least some worlds in which neither shaft is blocked. Since $d(i + \text{'}InA\text{'})$ selects those worlds in $i + \text{'}InA\text{'}$ in which shaft A is blocked, '$InA \Rightarrow \Box_d BlA$' is true at $\langle w, i\rangle$. Also, 'InA' is true at $\langle w, i\rangle$. But $d(i)$ is the nonempty subset of worlds in i in which neither shaft is blocked, so '$\Box_d BlA$' is false at $\langle w, i\rangle$. Thus, $\{\text{'}InA \Rightarrow \Box_d BlA\text{'}, \text{'}InA\text{'}\} \not\models_{Tr} \text{'}\Box_d BlA\text{'}$ and according to Kolodny and MacFarlane, this *modus ponens* inference is dangerous in certain hypothetical contexts.

5 Second Defense

However, this second attack also fails. Kolodny and MacFarlane, I will now argue, misdiagnose the problem with your practical deliberation. The trouble arises not at steps 5 and 8 where you apply *modus ponens* inside the two subproofs, but only at step 10 where you conclude the *proof by cases*.[36]

Consider the *informational background* of your deliberation. At first, your information leaves open the possibilities both that the miners are in shaft A and that they are in shaft B. However, in the hypothetical context triggered by your supposition that the miners are in shaft A, your provisional information rules out the possibility that the miners are in shaft B, so you do well to infer that you ought to block A, and therefore that either you ought to block A or you ought to block B. Likewise, in the hypothetical context triggered by your supposition that the miners are in shaft B, you do well to infer that this disjunction holds. But it does not follow from this that your original information is information on the basis of which either you ought to block A or you ought to block B. After all, a body of information that leaves open the possibilities both that the miners are in shaft A and that they are in shaft B certainly isn't a body of information that rules out the possibility that the miners are in one or the other shaft.

In a close variation of the miners example that also involves the use of *modus ponens* in a hypothetical context, you *can* conclude that either shaft A or shaft B ought to be blocked. Suppose that I am considering

[36] Though I do not discuss deontic modals in Bledin [2014], the diagnosis here will not surprise readers of my earlier work.

whether to block a shaft and I tell you that according to what I know either the miners *must* be in shaft A or they *must* be in shaft B. You can then reason as follows about what I ought to do with the sandbags on the basis of my information:

1	The miners must be in A or they must be in B	Premise
2	If the miners are in A, then A ought to be blocked	Premise
3	If the miners are in B, then B ought to be blocked	Premise
4	The miners must be in shaft A	Supposition
5	The miners are in shaft A	From 4
6	Shaft A ought to be blocked	From 2,5
7	Shaft A or shaft B ought to be blocked	From 6
8	The miners must be in shaft B	Supposition
9	The miners are in shaft B	From 8
10	Shaft B ought to be blocked	From 3,9
11	Shaft A or shaft B ought to be blocked	From 10
12	Shaft A or shaft B ought to be blocked	From 1,4-7,8-11

The main difference between this piece of deductive argumentation and the original one is that the first premise is now a disjunction of epistemic necessities.[37] However, unlike the earlier argumentation, your reasoning here is impeccable; you do well to advise me to take action and block one of the shafts. This suggests that the problem with the original miners example is not the use of *modus ponens.* To be sure, this second example also involves *proof by cases.* By attending to the informational background of your deliberation, though, one can explain why you can felicitously conclude that either shaft A or shaft B ought to be blocked in this second example but not in the first. In the second example, the salient body of information behind your inquiry is my knowledge state that either rules out the possibility that the miners are in shaft A or rules out the possibility that the miners are in shaft B. Suppose, without loss of generality, that my knowledge rules out that the miners are in A. In the hypothetical context triggered by your supposition that the miners

[37] This second example also involves inferences from $\ulcorner \Box_e \varphi \urcorner$ to φ. Kratzer [1991] and Veltman [1985] offer semantic theories on which this argument form is invalid. However, von Fintel and Gillies [2010] persuasively argue that '$\Box_e$' is strong.

must be in B, then, the salient information is just my knowledge state. Consequently, from the fact that either shaft A or shaft B ought to be blocked on the basis of the provisional information in each of the subproofs, it follows that either shaft A or shaft B ought to be blocked on the basis of my knowledge.

The take-home lesson from the above examples, I am suggesting, is not that *modus ponens* for the indicative conditional is unreliable in certain hypothetical contexts, but instead that we must exercise caution when using *proof by cases* in languages that contain informational modals and the indicative. Applying the fully general form of *proof by cases* in such languages can lead to absurd results.[38] But don't worry; I am not suggesting that we reject *proof by cases* outright. There are still various good forms of case-based reasoning and we should carefully catalogue these different argument forms.[39]

To complete my second defense, I should also mention that Kolodny and MacFarlane argue that *modus ponens* is objectionable in a different kind of environment—the hypothetical contexts of *reductio ad absurdum*. Suppose again that you do not know which shaft the miners are in. You might then reason as follows:

1	Neither shaft A nor shaft B ought to be blocked	Premise
2	If the miners are in A, then I ought to block A	Premise
3	Not: I ought to block A	From 1
4	The miners are in shaft A	Supposition
5	I ought to block A	From 2,4
6	$\perp$	From 3,5
7	The miners are not in shaft A	From 4-6

Given your evidence, you reflect that neither shaft ought to be blocked, but that if the miners are in A then you ought to block A. You then enter a hypothetical context by supposing that the miners are in shaft A. Applying *modus ponens* in this hypothetical context, you infer that you ought to block A. Recognizing that this conflicts with the first premise, you conclude by *reductio* that the miners are not in shaft A. However, your evidence leaves open the possibility that the miners are in shaft A. Something has gone horribly wrong.

[38] See Bledin [2014], §7, for another counterexample to the classical *proof by cases* involving epistemic rather than deontic modals.

[39] I make some headway on this project in my [2014].

Since *modus ponens* is the only "common factor" between this miners example and the original one, Kolodny and MacFarlane reject this rule. But is *modus ponens* really the source of the trouble in the *reductio* proof? It seems to me that Kolodny and MacFarlane again misdiagnose the problem by failing to appreciate the informational background of your practical deliberation. The trouble begins only at step 6 or 7, depending on how the supposition works at step 4. If your supposition consists in tentatively *adding* the information that the miners are in shaft A to the informational background of your deliberation, then you should not import part of the first premise into the subproof at step 6. After the supposition, your provisional information leaves open the possibility only that the miners are in shaft A, so you do well to infer that you ought to block A. But the first premise does not *persist* into the induced hypothetical context, so you should not derive the contradiction.[40]

On the other hand, your supposition might trigger a hypothetical context in which the provisional information rules out that the miners are not in shaft A but also retains the structural features corresponding to both premises.[41] In this case, you do well to derive the contradiction.

---

[40]Willer [2012] also diagnoses the problem with the *reductio* at step 6. He claims that importing part of the first premise into the subproof is objectionable because logical consequence is not right-monotonic. Specifically, Willer proposes the following "dynamic logical consequence" relation:

$<\varphi_1, ..., \varphi_n> \models_D \psi$ just in case there is no proper point $\langle w, i\rangle$ such that φ_1 is true at $\langle w, i\rangle$, φ_2 is true at $\langle w, i \boxplus \varphi_1\rangle$, ..., φ_n is true at $\langle w, i \boxplus ... \boxplus \varphi_{n-1}\rangle$, but ψ is false at $\langle w, i \boxplus ... \boxplus \varphi_n\rangle$

where $i \boxplus \varphi = i \cap \{w : \varphi$ is true at $\langle w, i\rangle\}$ (I use ordered sequences since $\models_D$ is sensitive to the order of premises). $<$ '$\neg\Box_d BlA \wedge \neg\Box_d BlB$', '$InA \Rightarrow \Box_d BlA$' $>\models_D$ '$\neg\Box_d BlA$' but $<$ '$\neg\Box_d BlA \wedge \neg\Box_d BlB$', '$InA \Rightarrow \Box_d BlA$', '$InA$' $>\not\models_D$ '$\neg\Box_d BlA$'. Thus, while you can infer in the main categorical context of your deliberation that it is not the case that you ought to block A, you cannot infer this in the hypothetical context triggered by the supposition that the miners are in shaft A.

Unlike Willer, I am suggesting that the problem with the *reductio* is at step 6 not because of the non-monotonicity of logical consequence—the relation $\models_I$ endorsed in §3 has standard structural properties—but because of how supposition affects the informational background of deliberation. Since a supposition can contract the space of live options under consideration, we must restrict what can be imported into subproofs (see Bledin [2014] for more details).

[41]In Bledin [2013], Appendix A, I distinguish this *lossless* kind of supposition from the earlier *lossy* kind. Lossy supposition triggers hypothetical contexts in which the premises of an argument can fail to hold. Lossless supposition always triggers hypothetical contexts in which one's provisional information incorporates everything that was incorporated before. (Full disclosure: I have since become less confident that a full theory of deductive inquiry appropriate to languages with informational modals and the indicative must acknowledge the non-standard lossless kind of supposition in addition to the standard lossy kind. But since I am still on the fence about this,

But while your information in the hypothetical context is degenerate—it can be explicated by the empty set $\emptyset$—you *cannot* conclude from this that the miners are not in shaft A. It follows only that your original evidence does not rule out that the miners are not in shaft A. That is, you are entitled to only the weaker conclusion that the miners *might* not be in shaft A.

6 Third Attack

So much for the second attack. Kolodny and MacFarlane have not given us reason to reject the thesis that "quasi-validity" coincides with deductively good argument. More particularly, they have not discredited the idea that *modus ponens* for the indicative conditional is a reliable rule of inference in both categorical and hypothetical deliberative contexts. Let me now turn to the final attack. Willer [2010] also argues, as follows, that *modus ponens* is objectionable in certain hypothetical contexts.[42]

Suppose that you find Sally rather cunning. In particular, you accept this Thomason conditional:

(P1) If Sally is lying, then I do not believe that she is lying.[43]

Further, suppose that you are a rational reflective agent. Not only do you fail to accept the Moore-paradoxical conjunction 'Sally is lying and I do not believe it'; in fact you accept its negation:

(P2) It is not the case that both Sally is lying and I do not believe it.

I leave open the possibility here that the trouble with the *reductio* arises at step 7.)

[42] As mentioned in n. 40, Willer [2012] defends *modus ponens* against Kolodny and MacFarlane. But his goal "is not to offer an unqualified defense of *modus ponens* but rather to contribute to the (no less interesting) project of determining the scope of its validity." (p. 449, n. 2)

[43] Willer himself employs the indicative conditional 'If Sally is deceiving me, then I do not believe it'. However, 'deceive' is arguably an *achievement* or *success verb*—the perlocutionary effect of deceit is the induction of inaccurate doxastic attitudes—so Willer's conditional is arguably an analytic truth. To avoid distracting complications stemming from this, I have substituted the non-success verb 'lie'.

You then reason as follows:

1	If Sally is lying, then I do not believe it	Premise
2	Not: Sally is lying and I do not believe it	Premise
3	Sally is lying	Supposition
4	I do not believe that Sally is lying	From 1,3
5	Sally is lying and I do not believe it	From 3,4
6	$\perp$	From 2,5
7	Sally is not lying	From 3-6

In the hypothetical context triggered by your supposition that Sally is lying, you infer by *modus ponens* that you do not believe that she is lying, and therefore that Sally is lying and you do not believe it. Recognizing that this conflicts with the second premise, you conclude the following:

(C) Sally is not lying.[44]

But this inference, Willer claims, is infelicitous:

> This is the wrong result. Certainly, clever women are not always loyal, so [you] should not be allowed to infer Sally's loyalty from her cleverness. (p. 298)

Willer concludes that *modus ponens* is logically invalid. Importantly, he does *not* conclude that this argument form is invalid because it fails to unrestrictedly preserve truth—Willer is sympathetic to the idea that indicative conditional declarative sentences are not truth-value bearers, and so semantic laws concerned with truth preservation do not apply to these sentences. He concludes rather that *modus ponens* is invalid because it licenses a bad inference.

[44]Willer has you infer directly from the first premise that either Sally is not lying or you do not believe it (on Willer's formulation of *modus ponens*, this inference rule licenses a transition from an indicative conditional to the corresponding material conditional), then infer directly from the second premise that either Sally is not lying or you believe it, and then finally conclude that Sally is not lying. I suspect that Willer avoids my *reductio* formulation to avoid the objection that in the hypothetical context where you suppose that Sally is lying you should no longer accept the first premise. However, this is not *my* objection and I find the *reductio* formulation more elegant.

7 Third Defense

Now, I think that Willer's attack also fails. To see why, we can expand our formal informational framework to handle belief reports. One can, I think, see the error of Willer's ways outside this framework—in short, he misjudges the probative force of a premise in his alleged counterexample by underestimating the gulf between not believing and belief in the negation. But the formalism will help sharpen my diagnosis.

Let 'H' abbreviate 'Sally is lying to Harry',[45] and let us add a belief operator 'Bel' to $\mathcal{L}$. To handle this new operator, our earlier definitions must be revised:

Def 11. A *model* $\mathcal{M} = \langle \mathcal{W}, \mathcal{V}, d, \mathcal{B} \rangle$ for $\mathcal{L}$ now consists of a nonempty set of worlds $\mathcal{W}$, an interpretation function $\mathcal{V}$, a deontic selection function d, and another function $\mathcal{B}$ mapping each world $w \in \mathcal{W}$ to an information state $i \in 2^{\mathcal{W}}$.[46] The set $\mathcal{B}(w) \subseteq \mathcal{W}$ explicates Harry's belief state in w.

Def 12. *Truth at a point* is defined by adding the following semantic clause to the clauses in Def 5 and Def 9 (cf. Yalcin [2011]):

$$\ulcorner Bel(\varphi) \urcorner \text{ is true at } \langle w, i \rangle \quad \text{iff} \quad \varphi \text{ is true at } \langle v, \mathcal{B}(w) \rangle \text{ for all } v \in \mathcal{B}(w)$$

The belief report 'Harry believes that Sally is lying to him' is true at $\langle w, i \rangle$ just in case $\mathcal{B}(w)$ incorporates the embedded sentence 'H'—that is, just in case Sally is lying to Harry in all of the ways the world might be that are left open by what Harry believes in w.[47]

With this formal semantics, we are well positioned to see where Willer goes wrong. Recall the first assumption: Harry believes that if Sally is lying to him then he does not believe that this is the case—formally, $\mathcal{B}(@) \rhd$ '$H \Rightarrow \neg Bel(H)$' where '@' designates the actual world.

[45] I have replaced the first-person pronoun 'me' with the proper name 'Harry' to sidestep distracting indexicality issues.

[46] Alternatively, one might cling to the standard relational semantics for modal languages by first introducing a binary accessibility relation $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$ between worlds and then letting $\mathcal{B}(w) = \{v : w\mathcal{R}v\}$.

[47] Modeling Harry's belief state as a single set of possible worlds is admittedly unrealistic. Given the semantic clause for belief reports, either Harry's beliefs are logically consistent and closed under logical consequence (when $\mathcal{B}(w) \neq \emptyset$) or he believes everything (when $\mathcal{B}(w) = \emptyset$). On a more realistic picture allowing for logical incoherence, Harry's doxastic state is *fragmented*. His total doxastic state might be modeled as a *subset* of $2^{\mathcal{W}}$ (Stalnaker [1984]) or as a function from partitions to subpartitions of $\mathcal{W}$ (Yalcin [2011]). We might also place a quantitative measure over Harry's doxastic alternatives and evaluative belief reports relative to this measure. However, for present purposes, this more complicated structure is unnecessary; the simplest, most idealized model and the semantic clause for 'Bel' in Def 12 will do.

The second assumption is that Harry is rational and reflective.[48] In our possible-worlds framework, Harry's belief state $\mathcal{B}(w)$ in each world $w \in \mathcal{W}$ is nonempty and has the following property:

Def 13. Belief state $\mathcal{B}(w)$ is *reflective* iff $\mathcal{B}(v) = \mathcal{B}(w)$ for all $v \in \mathcal{B}(w)$.[49]

If $\mathcal{B}(w)$ is reflective, then it is easy to check that for all $\varphi \in S_{\mathcal{L}}$, $\mathcal{B}(w) \rhd \varphi$ iff $\mathcal{B}(w) \rhd \ulcorner Bel(\varphi) \urcorner$.[50] So if Harry believes that such and such, then he believes that he believes it. Moreover, if $\mathcal{B}(w)$ is both reflective and nonempty, then for all $\varphi \in S_{\mathcal{L}}$, $\mathcal{B}(w) \not\rhd \varphi$ iff $\mathcal{B}(w) \rhd \ulcorner \neg Bel(\varphi) \urcorner$.[51] So if Harry does not believe that such and such, then he believes that he does not believe it.

Because Harry is a rational reflective agent, Willer thinks that Harry must also believe that it is not the case that Sally is lying and he does not believe it—formally, $\mathcal{B}(@) \rhd$ '$\neg(H \wedge \neg Bel(H))$':

> Moore paradoxical constructions are unacceptable to agents who are rational. Unacceptability comes in different flavors. The agent might be unable to accept φ since he lacks sufficient evidence in support of φ. The agent might also be unable to accept φ since φ is a priori absurd. For instance, φ might be an obvious contradiction. In such cases, not only is φ unacceptable, but also is the agent rationally committed to accept $\neg\varphi$. Neither accepting φ nor $\neg\varphi$ is not an option: Unacceptability of φ commits

[48] According to Willer, reflective agents have perfect higher-order knowledge of what they believe and do not believe:

> We want agents to be reflective, their belief sets encoding not only first-order beliefs but also being closed under what the agent considers to be his own doxastic state. (p. 295)

Willer himself cashes this out in a syntactic framework where belief states are sets of sentences but, as I show, reflectivity can also be modeled in a possible-worlds framework.

[49] 'Reflective' as used here is a technical term applicable to the mathematical objects explicating what Harry believes, not an informal term applicable to Harry himself. But the technical and nontechnical uses are related: Harry is reflective in w iff $\mathcal{B}(w)$ is reflective.

On the standard relational semantics (see n. 46), the reflectivity requirement amounts to the requirement that $\mathcal{R}$ is *transitive* $(\forall w, v, u((w\mathcal{R}v \wedge v\mathcal{R}u) \supset w\mathcal{R}u))$ and *Euclidean* $(\forall w, v, u((w\mathcal{R}v \wedge w\mathcal{R}u) \supset v\mathcal{R}u))$. The requirement that $\mathcal{B}(w)$ is nonempty amounts to the requirement that $\mathcal{R}$ is *serial* $(\forall w \exists v(w\mathcal{R}v))$.

[50] This equivalence holds when $\mathcal{B}(w) = \emptyset$ since $\emptyset \rhd \varphi$ for all $\varphi \in S_{\mathcal{L}}$. In general, $\mathcal{B}(w) \rhd \varphi$ iff $\mathcal{B}(v) \rhd \varphi$ for all $v \in \mathcal{B}(w)$ iff $\ulcorner Bel(\varphi) \urcorner$ is true at $\langle v, \mathcal{B}(w) \rangle$ for all $v \in \mathcal{B}(w)$ iff $\mathcal{B}(w) \rhd \ulcorner Bel(\varphi) \urcorner$.

[51] This equivalence fails to hold when $\mathcal{B}(w) = \emptyset$ since $\emptyset \rhd \ulcorner \varphi \wedge \neg Bel(\varphi) \urcorner$. When $\mathcal{B}(w) \neq \emptyset$, $\mathcal{B}(w) \not\rhd \varphi$ iff $\mathcal{B}(v) \not\rhd \varphi$ for all $v \in \mathcal{B}(w)$ iff $\ulcorner \neg Bel(\varphi) \urcorner$ is true at $\langle v, \mathcal{B}(w) \rangle$ for all $v \in \mathcal{B}(w)$ iff $\mathcal{B}(w) \rhd \ulcorner \neg Bel(\varphi) \urcorner$.

> any rational agent to acceptance of $\neg\varphi$. Moorean paradoxical constructions are unacceptable in the latter sense: It is simply absurd for a rational agent to judge true both that φ holds and that he does not believe that φ holds. (p. 298)

But this, I submit, is Willer's error. I will happily grant that a rational reflective agent cannot accept Moore paradoxical constructions. Since $\mathcal{B}(@)$ is nonempty and reflective, for instance, $\mathcal{B}(@) \not\rhd$ '$H \wedge \neg Bel(H)$'.[52] However, I will not grant that a rational reflective agent must accept the negations of these paradoxical constructions. Information states that do not incorporate φ will incorporate $\ulcorner \Diamond_e \neg\varphi \urcorner$, but $\rhd \ulcorner \neg\varphi \urcorner$ is typically a far stronger constraint than $\rhd \ulcorner \Diamond_e \neg\varphi \urcorner$. If $\mathcal{B}(@) \rhd$ '$\neg(H \wedge \neg Bel(H))$' and $\mathcal{B}(@)$ is reflective, then either $\mathcal{B}(@) \rhd$ '$\neg H$' or $\mathcal{B}(@) \rhd$ 'H'—that is, either Harry's belief state rules out that Sally is lying, or it rules out that Sally is not lying. Surely rationality and reflectivity alone cannot mandate that Harry's belief state have this structure.

Indeed, if Harry's belief state *has* this structure, then he can infer that Sally is not lying to him. Putting the initial two assumptions together, his belief state cannot rule out the possibility that Sally is not lying: if $\mathcal{B}(@) \rhd$ 'H', then $\mathcal{B}(@) +$ 'H' $= \mathcal{B}(@)$ and $\mathcal{B}(@) \rhd$ '$Bel(H)$' by reflectivity, so contrary to the first assumption $\mathcal{B}(@) \not\rhd$ '$H \Rightarrow \neg Bel(H)$'. Since $\mathcal{B}(@) \not\rhd$ 'H', $\mathcal{B}(@) \rhd$ '$\neg H$'. A nonempty reflective belief state that incorporates the two premises of Willer's argument is therefore also an information state that rules out the possibility that Sally is lying to Harry. To say, as Willer does, that Harry mistakenly infers Sally's loyalty from her cleverness is to ignore the crucial role of the second premise in the argument.

Even if $\mathcal{B}(@)$ is *not* a reflective state, Harry can argue from the premises '$H \Rightarrow \neg Bel(H)$' and '$\neg(H \wedge \neg Bel(H))$' to the conclusion '$\neg H$'. If $\mathcal{B}(@) \rhd$ '$H \Rightarrow \neg Bel(H)$', then Harry's belief state rules out that Sally is lying and he believes it. If $\mathcal{B}(@) \rhd$ '$\neg(H \wedge \neg Bel(H))$', then his belief state also rules out that Sally is lying and he does not believe it. Consequently, *any* belief state that incorporates both premises of the argument must incorporate the conclusion that Sally is not lying to Harry.

Willer's argument, I conclude, does not establish that *modus ponens* for the indicative conditional is unreliable. The application of this rule does not lead Harry to an unwarranted conclusion. In fact, I think we

---

[52] Suppose that $\mathcal{B}(@) \rhd$ '$H \wedge \neg Bel(H)$'. Since $\mathcal{B}(@) \rhd$ 'H' and $\mathcal{B}(@)$ is reflective, $\mathcal{B}(@) \rhd$ '$Bel(H)$'. But then $\mathcal{B}(@) \rhd$ '$Bel(H) \wedge \neg Bel(H)$', contradicting $\mathcal{B}(@) \neq \emptyset$. Similar reasoning establishes that $\mathcal{B}(@) \not\rhd$ '$Bel(H) \wedge \neg H$', $\mathcal{B}(@) \not\rhd$ '$H \wedge \Diamond_e \neg Bel(H)$', $\mathcal{B}(@) \not\rhd$ '$H \wedge Bel(\Diamond_e \neg H)$', and so forth.

can say a bit more: not only is the argument from '$H \Rightarrow \neg Bel(H)$' and '$\neg(H \wedge \neg Bel(H))$' to '$\neg H$' a good argument, but the *reductio* in §6 is *good argumentation*. Consider the informational background of this episode of deliberation. When Harry supposes that Sally is lying, he thereby enters a hypothetical context where the open possibilities are all ones in which Sally is lying. Since Harry's provisional information still rules out the possibility that Sally is lying and he believes it, he does well to recognize that these live possibilities are also ones where he does not believe that Sally is lying. However, Harry's information also rules out the possibility that Sally is lying and he does not believe it. There are in fact *no* open possibilities where Sally is lying—Harry's degenerate information in the hypothetical context rules out everything. Thus, he does well to conclude that Sally is loyal.[53]

8 Conclusion

So much for the third and final attack. Thomason conditionals and Moore's Paradox do not undermine *modus ponens* either. We have not yet encountered a *bona fide* counterexample to this rule. There might, of course, be other lines of attack besides the three I discuss in this paper, but I do not know of them. In any case, it should be clear from the preceding discussion that I do not think any further attacks will succeed. In the course of defending *modus ponens* against McGee, Kolodny and MacFarlane, and Willer, I have put forward a semantics for the indicative conditional and an informational conception of validity on which *modus ponens* is valid. If this theory is more-or-less correct, then no counterexamples to *modus ponens* are forthcoming. Far from leading reasoners astray, this inference rule can help establish what is so according to information that incorporates the premises of an argument.

---

[53]On the simple model I have been working with, Harry *already* believes that Sally is loyal before he engages in deductive inquiry. To make sense of how *reductio* can be a learning experience in which Harry acquires new beliefs, we might model his belief state $\mathcal{B}(@)$ as a set of fragments $\{\mathcal{B}(@)_n\}_{n\in\mathbb{N}}$ (recall n. 47) and upgrade the semantic clause for belief reports accordingly:

$$\ulcorner Bel(\varphi)\urcorner \text{ is true at } \langle w,i\rangle \quad \text{iff} \quad \mathcal{B}(w)_n \rhd \varphi \text{ for some } n \in \mathbb{N}$$

Prior to Harry's deliberation, the story might now go, $\mathcal{B}(@) = \{\mathcal{B}(@)_1, \mathcal{B}(@)_2\}$, where $\mathcal{B}(@)_1 \rhd$ '$H \Rightarrow \neg Bel(H)$' and $\mathcal{B}(@)_2 \rhd$ '$\neg(H \wedge \neg Bel(H))$', but $\mathcal{B}(@)_1 \not\rhd$ '$\neg H$' and $\mathcal{B}(@)_2 \not\rhd$ '$\neg H$'. The *reductio* removes this fragmentation. After deliberation, $\mathcal{B}(@) = \{\mathcal{B}(@)_3\}$ where $\mathcal{B}(@)_3 \rhd$ '$H \Rightarrow \neg Bel(H)$', $\mathcal{B}(@)_3 \rhd$ '$\neg(H \wedge \neg Bel(H))$', and $\mathcal{B}(@)_3 \rhd$ '$\neg H$'.

References

Ernest W. Adams. *The Logic of Conditionals*. Reidel, Dordrecht, 1975.

Alan Ross Anderson and Nuel Belnap Jr. *Entailment, Volume I*. Princeton University Press, Princeton, 1975.

Jon Barwise and John Etchemendy. *Language, Proof and Logic*. CSLI Publications, Stanford, 1999.

J. C. Beall. *Spandrels of Truth*. Oxford University Press, Oxford, 2009.

Garrett Birkhoff and John Von Neumann. The logic of quantum mechanics. *Annals of Mathematics*, 37(4):823–843, 1936.

Justin Bledin. *Logic Informed*. PhD thesis, University of California, Berkeley, 2013.

Justin Bledin. Logic informed. *Mind*, 123(490):277–316, 2014.

Rachel Briggs. Interventionist counterfactuals. *Philosophical Studies*, 160(1):139–166, 2012.

John Broome. Normative requirements. *Ratio*, 12:398–419, 1999.

L. E. J. Brouwer. Intuïtionistische verzamelingsleer. *KNAW Verslagen*, 29, 1921. Translated as 'Intuitionist Set Theory' in Paolo Mancosu, editor, *From Brouwer to Hilbert*, pages 23–27. Oxford University Press, 1998.

L. E. J. Brouwer. Intuïtionistische splitsing van mathematische grondbegrippen. *KNAW Verslagen*, 32, 1923. Translated as 'Intuitionist Splitting of the Fundamental Notions of Mathematics' in Paolo Mancosu, editor, *From Brouwer to Hilbert*, pages 286–289. Oxford University Press, 1998.

Fred Dretske. The case against closure. In Matthias Steup and Ernest Sosa, editors, *Contemporary Debates in Epistemology*, pages 13–25. Blackwell, 2005.

Dorothy Edgington. Do conditionals have truth conditions? *Critica: Revista Hispanoamericana de Filosofia*, 18(52):3–39, 1986.

Hartry Field. Truth and the unprovability of consistency. *Mind*, 115 (459):567–606, 2006.

Hartry Field. *Saving Truth From Paradox*. Oxford University Press, Oxford, 2008.

Hartry Field. What is the normative role of logic? *Proceedings of the Aristotelian Society Supplementary Volume*, LXXXIII:251–268, 2009a.

Hartry Field. Pluralism in logic. *Review of Symbolic Logic*, 2:342–359, 2009b.

Hartry Field. What is logical validity? In Colin Caret and Ole Hjortland, editors, *Foundations of Logical Consequence*. Oxford University Press, forthcoming.

Joseph S. Fulda. Vann McGee's counterexample to Modus Ponens: An enthymeme. *Journal of Pragmatics*, 42:271–273, 2010.

Anthony Gillies. Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616, 2004.

Gilbert Harman. *Change in View: Principles of Reasoning*. MIT Press, Cambridge, 1986.

David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes From Kaplan*, pages 481–563. Oxford University Press, 1989.

Bernard Katz. On a supposed counterexample to Modus Ponens. *Journal of Philosophy*, 96(8):404–415, 1999.

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

Angelika Kratzer. What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, 1(3):337–355, 1977.

Angelika Kratzer. Modality. In Armin von Stechow and Dieter Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, pages 639–650. de Gruyter, 1991.

Henry Kyburg. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, 1961.

David Lewis. Adverbs of quantification. In E. L. Keenan, editor, *Semantics of Natural Language*, pages 3–15. Cambridge University Press, 1975.

John MacFarlane. In what sense (if any) is logic normative for thought? Unpublished manuscript.

David C. Makinson. The paradox of the preface. *Analysis*, 25(6):205–207, 1965.

Vann McGee. A counterexample to Modus Ponens. *Journal of Philosophy*, 82(9):462–471, 1985.

Christian Piller. Vann McGee's counterexample to Modus Ponens. *Philosophical Studies*, 82:27–54, 1996.

Graham Priest. The logic of paradox. *Journal of Philosophical Logic*, 8: 219–241, 1979.

Hilary Putnam. Is logic empirical? In Robert Cohen and Marx Wartofsky, editors, *Boston Studies in the Philosophy of Science*, volume 5, pages 216–241. Reidel, 1968.

Walter Sinnott-Armstrong, James Moor, and Robert Fogelin. A defense of Modus Ponens. *Journal of Philosophy*, 83(5):296–300, 1986.

Robert Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Oxford University Press, 1968.

Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975.

Robert Stalnaker. *Inquiry*. MIT Press, Cambridge, 1984.

Alfred Tarski. O pojciu wynikania logicznego. *Przeglad Filozoficzny*, 39, 1936a. Translated as 'On the Concept of Logical Consequence' in *Logic, Semantics, Metamathematics*, pp. 409–420, 1956.

Alfred Tarski. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. Ed. and Trans. J. H. Woodger, Clarendon Press, Oxford, 1956.

Jan Łukasiewicz. Philosophische bermekungen zu mehrwertigen systemen des aussagenkalküls. *Comptes rendus des seances de la Societe des Sciences et des Lettre de Varsovie*, 23, 1930. Translated as 'Philosophical Remarks on Many-Valued Systems of Propositional Logic' in *Selected Works*, pp. 153–178, 1970.

Jan Łukasiewicz. *Selected Works*. North Holland, Amsterdam, 1970.

Bas van Fraassen. Review of Brian Ellis, *Rational Belief Systems*. *Canadian Journal of Philosophy*, 10(3):497–511, 1980.

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

Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.

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

Malte Willer. New surprises for the Ramsey test. *Synthese*, 176(2):291–309, 2010.

Malte Willer. A remark on iffy oughts. *Journal of Philosophy*, 109(7):449–461, 2012.

Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, 2007.

Seth Yalcin. Nonfactualism about epistemic modality. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 295–332. Oxford University Press, 2011.