Modus Ponens and the Logic of Decision

Nate Charlow

1 Introduction

Some time in 2013 I came to accept, for very mundane reasons, that the best thing for me was to quit smoking. Though I enjoyed smoking, I didn't want to get cancer from it.[1] My mundane reasons were grounded in equally mundane facts about my preferences. Although I certainly preferred the outcome in which I smoked but did not get cancer to the one in which I did not smoke and did not get cancer, I vastly preferred the latter outcome to the outcome in which I smoked and got cancer. (Worst of all possible outcomes was the one in which I did not smoke and still got cancer.) I take it as obvious that I was *right* then. Smoking greatly increases your risk of cancer, and this is a risk you should obviously avoid, if you have preferences like mine.

1.1 The Puzzle

Little did I know then, there was an argument to a very different (and, it should be stressed, very implausible) conclusion, but with an apparently quite strong claim to logical soundness. Consider this accurate, if abstract, representation of the "decision problem" I faced then.[2]

	CANCER	¬CANCER
SMOKE	☹	☺☺
DON'T	☹☹	☺

Given (we might say against) this representation of my decision, statements of conditional preference (1a) and (1b) are easily heard as true: they respectively summarize the information in the CANCER and ¬CANCER columns of the above decision table. (1c) is easily heard as true as well, and seems to follow from (1a) and (1b). But, by apparent application of modus ponens, I arrive at something unacceptable—that I cannot hear as true—that is, (1d).

(1) a. If I get cancer, it's better to smoke. (≈smoking is better than not smoking)
b. If I don't get cancer, it's better to smoke.
c. So, if I get cancer or I don't, it's better to smoke.
d. #So, it's better to smoke.

I said (1a) and (1b) are heard as true against this representation of the decision, but maybe you are doubtful. If so, apply the Ramsey Test. Suppose you have preferences like mine. Now suppose that you'll get cancer. Which do you prefer: smoking or not smoking? Evidently, smoking. Bearing this in mind, you should accept (1a). Suppose that you won't get cancer. Which do you prefer: smoking or not smoking? Evidently, smoking. Bearing this in mind, you should accept (1b). And (1c)—more precisely, an intended reading of (1c)—is true, and appears to follow from (1a) and (1b).[3]

---

[1] Throughout I use "cancer" as a metonym for any form of cancer causally linked to smoking tobacco.

[2] I rely for now on an intuitive notion of what a decision problem is: a decision problem Π is something that determines a preference ordering over outcomes, where o is an outcome in Π iff o is the conjunction $a \wedge s$, for some action a available in Π and some state s that is relevant in Π. In this decision problem, the available actions are the agent smoking or not smoking; the relevant states are the agent getting cancer or not getting cancer. For a refinement of this intuitive notion, see §4.1.

[3] A note to decision theorists: I realize the decision is ill-formed, by any metric of well-formedness you might choose. But that does not alter the apparent logical facts: if modus ponens is valid *and (1c) is true*, it's better to smoke, which is contrary to fact.

A note to linguists: I realize that the logical representations I have chosen for (1a)–(1c) ignore matters like tense and aspect—matters of possible significance for evaluating statements of (conditional) preference like these. But unless the representation of tense and aspect renders the move from (1c) to (1d) a non-instance of modus ponens or renders $(p \vee \neg p)$ a non-instance of $\top$—and I cannot see how to make out a story on which either would be the case—tense and aspect are irrelevant to this problem, and we do well to abstract away from the complications would accompany their introduction.

1.2 *Logical Preliminaries*

The reasoning in (1) makes use of two inference rules: (CA), a fairly standard principle of conditional logic, and modus ponens (henceforth MP).

$$
\begin{array}{|ll}
p \Rightarrow r & \\
q \Rightarrow r & \\
(p \vee q) \Rightarrow r & \text{CA}
\end{array}
$$

(CA) is assumed by most theorists of the conditional (including both Lewis and Stalnaker), but is rejected in certain context-shifting and dynamic frameworks (a point to which we will return below). But, for present purposes, I believe it is possible simply to bracket the question of its general validity. Given (1a) and (1b), (1c) is read *as true*; on the intended reading, it *summarizes* the information contained in (1a) and (1b).[4] We can also bracket the question of whether (1c) has a false reading (as seems to me very likely). It seems clearly to have a true reading, on which it is summarizing the information expressed in (1a) and (1b). That reading—hereafter the "Intended Reading"—is the reading at issue here.

Notice that, given (MP), (CA) allows us to derive a standard elimination rule for disjunction, abbreviated (∨E).

$$
\begin{array}{|ll}
p \vee q & \\
p \Rightarrow r & \\
q \Rightarrow r & \\
r & \vee\text{E}
\end{array}
\qquad\qquad
\begin{array}{|ll}
p \vee q & \\
p \Rightarrow r & \\
q \Rightarrow r & \\
(p \vee q) \Rightarrow r & \text{CA} \\
r & \text{MP}
\end{array}
$$

The availability of (∨E) would allow us to run a different argument for smoking:[5]

(2) a. If I get cancer, it's better to smoke.
b. If I don't get cancer, it's better to smoke.
c. #So, it's better to smoke.

A (very tentative) hypothesis is that both (1) and (2) rest ultimately on applications of (MP) that are, for some reason, illicit.[6] (CA), whether or not it is a generally appropriate rule governing the logic of the conditional, appears to preserve truth in the smoking scenario: on the Intended Reading, (1c) is acceptable, given (1a) and (1b). (MP) is, it would seem, the only remaining culprit.

And a note to Aristotelians: while it may be tempting to deny Excluded Middle for future contingents, notice that intuitive failures of modus ponens, like (19), need not utilize future contingents. Notice also that many intuitively valid inferences (e.g., the conclusion of a valid dominance argument) seem to rely on the validity of Excluded Middle for future contingents.

Finally, if you are tempted to respond to this case by appealing to the thesis that $(p \vee \neg p)$ is ambiguous between an alternative (or inquisitive) semantic value and a classical Boolean semantic value, I discuss this possibility in detail in §3.

[4] Though I do not want to rest anything on this claim, the following is a plausible constraint on the logic of the conditional: $(p \Rightarrow r) \wedge (q \Rightarrow r) \dashv\vdash (p \vee q) \Rightarrow r$. The left-to-right direction encodes (CA). The right-to-left direction encodes Simplication of Disjunctive Antecedents (SDA). For more on SDA, see §3.

[5] The argument in (2) is an example of a classic fallacious dominance argument. Fallacious dominance arguments similar to (2) are the focus of Cantwell (2006), and a focus Gibbard & Harper (1981). Although neither Cantwell nor Gibbard and Harper center modus ponens in their discussions, as I do here, we are all interested in the same broad target: a logical account of why intuitively good dominance reasoning is sound, and intuitively bad dominance reasoning is unsound.

[6] The suggestion here is that (2) relies tacitly on (MP), since (∨E) is established by appeal to more basic principles within a proof theory for the conditional, namely, (CA) and (MP).

1.3 Stalnaker on Fatalism

Stalnaker (1975) discusses Dummett (1964)'s intuitively similar argument for "fatalism."[7]

> The setting of [Dummett's] example is wartime Britain during an air raid. I reason as follows: "Either I will be killed in this raid or I will not be killed. Suppose that I will. Then even if I take precautions I will be killed, so any precautions I take will be ineffective. But suppose I am not going to be killed. Then I won't be killed even if I neglect all precautions; so, on this assumption, no precautions are necessary to avoid being killed. Either way, any precautions I take will be either ineffective or unnecessary, and so pointless." (280)

Stalnaker's diagnosis is this:

> [T]he conclusion follows *validly* from the premiss, provided that the sub-arguments are *valid*. But it is not correct that the conclusion is a *reasonable inference* from the premiss, provided that the sub-arguments are *reasonable inferences* [...T]he sub-arguments are reasonable, but not valid, and this is why the argument fails [i.e., to be valid or reasonable]. (281)

I take Stalnaker to be suggesting that conditionals (3) and (4) are supported or established in this case *by conditional proof* (i.e. derivation of the consequent under supposition of the antecedent).

(3) If I am killed in the air raid, then precautions are pointless.

(4) If I am not killed in the air raid, then precautions are pointless.

Here, then, is a natural representation of the reasoning at issue in this case:

| I am killed in the air raid or I am not.
| | Suppose I am killed in the air raid.
| |---
| | Then, precautions are pointless.
| So, if I am killed in the air raid, then precautions are pointless.
| | Suppose I am not killed in the air raid.
| |---
| | Then, precautions are pointless.
| So, if I am not killed in the air raid, then precautions are pointless.
| So, precautions are pointless.

The difficulty with this reasoning, Stalnaker observes, is that, in each sub-derivation, the consequent cannot strictly be *derived* under the relevant supposition: it is certainly not a *logical* consequence of your being killed in the air raid that precautions are pointless. It is, Stalnaker allows, *reasonable to infer* that precautions are pointless, on the supposition that you are killed in the air raid. But (∨E) is not a rule that characterizes reasonable inference, as this example well-illustrates: from (3) and (4) it is not reasonable to infer that precautions are pointless. The fatalist's "proof" is, therefore, both *invalid* (since the sub-arguments establishing (3) and (4) are invalid) and *unreasonable* (since (∨E) is not a rule that characterizes reasonable inference). It is a failure, twice over.

But Stalnaker's diagnosis of Dummett's argument for fatalism does not really help with (1). Notice that the conditionals in (1) were provided directly by context, not established or supported by conditional proof: (1a) and (1b) are, very simply, *true* in their context of use. Since Stalnaker endorses (CA) and

[7] Dreier (2009) also discusses fatalistic (∨E) arguments, and concurs that they illustrate failures of (MP), on the grounds that "In some contexts [∨E] is suspect, but not here, I take it" (127–8). (Thanks to [redacted] for this comparison.) As far as I am aware, Dreier and I are the only two to diagnose fatalistic reasoning in this way (though of course Kolodny & MacFarlane (2010) diagnose apparent failures of Constructive Dilemma of the sort apparently attested in their Miner Case as failures of (MP) as well). I take it, however, that the move that Dreier makes in diagnosing this case, though correct, is not presently dialectically available, for two reasons. First, (∨E) *is* generally regarded as the suspect principle in cases of this general type (e.g., the Miner Case) (Willer 2012; Yalcin 2012b; Bledin 2014). Second, if indeed the logical status of (∨E) depends on the logical status of (MP) (as suggested in §1.2), giving up (MP) means giving up (∨E), even for arguments in which (MP) is not explicitly invoked, like (2).

(MP), and since Stalnaker is committed to the falsity of (1d), he is committed to denying that both (and presumably either) of (1a) and (1b) have the (true/acceptable) readings they seem plainly to have.[8] I take it that it would be preferable to offer some account of these plain readings, rather than to dismiss their semantic possibility a priori.[9]

1.4 Plan

Challenges to the validity of modus ponens (Kolodny & MacFarlane 2010) and modus tollens (Yalcin 2012b) have been the subject of intense interest in recent work in the formal semantics of natural language. There is presently some agreement in the literature that Kolodny and MacFarlane's challenge can be avoided, if the notion of logical consequence is understood aright (Willer 2012; Yalcin 2012b; Bledin 2014). The viability of Yalcin's counterexample to modus tollens has meanwhile been challenged on the grounds that it fails to take the phenomenon of contextual domain restriction (and context-sensitivity more generally) into proper account (Stojnić 2017; Schulz 2018).

This paper will describe a handful of cases in the mold of (1) that I will informally call "counterexamples" to modus ponens, and it will show that strategies developed for handling extant challenges to modus ponens and modus tollens do not account for these cases.[10] My primary aim here will be to argue that this "counterexample" (really class of "counterexamples") presents a genuine theoretical puzzle, one addressed neither by easy appeal to context-sensitivity nor by appeal to a novel—whether Dynamic or Informational or Inquisitive—account of the relation of logical consequence.

This paper's other aim will be to pitch a positive account of these cases (and to situate theorizing about their logic and semantics more broadly in the literature on formal models of rational choice and belief). On the account I propose, judgments of *minimal acceptability* (relative to a context of use) are not closed under modus ponens (although judgments of what I will call *reasonable acceptability*, relative to a context of use, are). In a case like (1), the contextually *salient* decision state plainly satisfies (1c) (and so (1c) is minimally acceptable in that context). But that decision state plainly does not satisfy (1d), and so (1d) is not minimally acceptable in that context (in which event, the minimal acceptability of the premises of a modus ponens argument does not generally imply the minimal acceptability of its conclusion). What

[8] Cantwell (2006) argues that, given Stalnaker's (similarity- or nearness-based) understanding of a conditional notion like 'the value if I do *a*', there is (independent) reason to regard (sentences like) (1a) and (1b) as false. On a similarity-based account, the values realizable if I do *a* are those witnessed at the nearest *a*-worlds. Since smoking *makes near* worlds where you get cancer, the set of values realizable if you smoke includes some very low values indeed. This is, Cantwell suggests, sufficient to falsify 'smoking is better than not smoking' in the relevant context: the values realizable if you smoke do not (generally) exceed the values if you don't. Assuming modus ponens, it follows that at least one of (1a) or (1b) must be false. Although this is the theory of cases like (1) that is generated by a Stalnakerian theory, and the theory is consistent, this obviously cannot be taken as evidence *for* the Stalnakerian theory of such cases. In fact, it is prima facie evidence *against* such a theory, as it suggests that the theory is insufficiently expressive. On the relevant readings of (1a) and (1b), it is possible to accept both while denying (1d). The Stalnakerian theory says that accepting (1a) and (1b) while denying (1d) is straightforwardly logically inconsistent, which is apparently contrary to fact. (All this said, it should be noted that Cantwell shares the sense that the badness of spurious dominance arguments traces to failures of independence, although we pursue this sense in different theoretical frameworks.)

[9] This presents a contrast with Dummett's argument, where it is easier to dismiss (3) and (4) a priori. Reasonable precautions always have *a point*, namely, *risk-reduction*; compare, e.g., Ayer (1964) and the statement from Hospers (1967) quoted at Buller (1995: 112). (In particular, in the event that you are not killed in the air raid, you would typically conclude that there is some chance that taking precautions *saved your life* by reducing your risk and so had a clear point after all.) Since the truth of (3) and (4) cannot be established directly (i.e., by appeal to semantic intuition), an indirect (e.g., logical) argument for their truth is required; but the obvious indirect argument (by conditional proof) is a logical failure, or so Stalnaker argues.

A reader might reasonably wonder if I my own argument is caught up in something like this dialectic: have I not also argued for (1a) and (1b) using something like conditional proof? In reply: I have, in fact, been careful to avoid this. Yes, I have tried to cajole the reader into accessing the relevant readings of (1a) and (1b) by suggesting the Ramsey Test. But this is not strictly intended as an *argument*, indirect or otherwise, in favor of accepting (1a) and (1b); it is rather an *instruction* for how to access the intended—and clearly true (I claim)—readings of (1a) and (1b), i.e., via application of the Ramsey Test.

A reader might also reasonably wonder about the prospects for reviving the strategy of Ayer and Hospers, i.e., about denying that (1a) and (1b) are true; perhaps it is *always* better to avoid the risk of developing cancer associated with smoking, in which case it is *always* better not to smoke, in which case, no conditional of the form 'if ϕ, then it is better to smoke' could be regarded as true. This is akin to the objection from Subjectivism considered in Kolodny & MacFarlane (2010). As they note, a clear difficulty with Subjectivism is empirical: it is a serious cost to render conditionals like (1a) and (1b) false, when a preponderance of competent speakers seem able to access a true reading. See Kolodny & MacFarlane (2010: §1.2) for another argument against Subjectivism.

[10] There seems to be, in other words, a more general phenomenon at work—something readers should bear in mind before attempting to develop one-off treatments of any of the paper's individual examples.

I will call the *reasonable acceptability* of the premises of a modus ponens argument *does*, however, imply the reasonable acceptability of its conclusion. But (1c) cannot be reasonably accepted relative to the contextually salient decision state (since the contextually salient decision state violates act-state independence—it is, in a for-now informal sense, *malformed*). I will also pitch a general understanding of reasonability, according to which reasonable states of decision are states which are not subject to proof-theoretic, or "narrowly" logical, commitments that are unacceptable from the decision state's own point of view.

2 Modus Tollens

Yalcin (2012b) described what he called a counterexample to modus tollens (MT).[11]

$$\begin{array}{|ll} p \Rightarrow q & \\ \neg q & \\ \neg p & \text{MT} \end{array}$$

In Yalcin's counterexample, there is a bag of marbles, varying in color and size as follows, from which one is drawn and concealed from you.

	BLUE	RED
BIG	10	30
SMALL	50	10

Sentences (5a) and (5b) are judged acceptable in this scenario. But, by application of (MT), we arrive at something unacceptable: (5c).

(5) a. If it's big it's likely red.
b. It's not likely red.
c. #So, it's not big.

$$\begin{array}{r|ll} & p \Rightarrow \triangle q & \\ & \neg\triangle q & \\ \# & \neg p & \text{MT} \end{array}$$

This section describes two responses to Yalcin's apparent counterexample to (MT). On both, the questionable inference is actually *not an instance of* (MT). On one, this is because the logical forms suggested for the premises of (5) are incorrect (because under-described). On another, this is because (MT) is claimed to be a rule governing preservation of truth with respect to a single context, and in (5) there is an illicit context-shift. I am not primarily interested in whether these replies blunt the force of Yalcin's counterexample (though the second reply is, in my view, quite powerful). My aim in this section is to establish that they do not extend to the counterexample to (MP) presented above.

2.1 Domain Restriction

It is commonly thought that 'if'-clauses semantically function to introduce *domain restrictions* for downstream quantificational expressions (in particular: modal, preferential, probabilistic, or epistemic operators).[12] Letting $\triangle$ be any such operator, the idea is that the relevant operators are binary, i.e., have as arguments a restriction and nuclear scope:

$\triangle_{\text{RESTRICTOR}}$ SCOPE

[11]This section is fairly technical, but easily skipped (especially if you are already satisfied of the truth of (1c) or uninterested in issues surrounding Yalcin's alleged counterexample to MT). In summary, this section argues that none of the strategies for dealing with Yalcin's apparent counterexample to (MT) supply any explanation of the apparent error in going from (1c) to (1d).

[12]The notion that 'if'-clauses are restrictors for downstream quantifiers has been developed extensively by Angelika Kratzer (see a.o. Kratzer 1981, 1991, 2012). The implementation here is not quite Kratzer's, since she treats the semantics of the 'if'-clause as *exhausted* by its restrictive function. I consider Kratzer's actual analysis of the relevant conditionals below.

The antecedent of the conditional supplies the operator's restriction argument.

$$p \Rightarrow \triangle q \;:=\; p \Rightarrow \triangle_p q$$

The suggestion (drawn from Stojnić 2017; Schulz 2018) is that, once we disambiguate (5) along these lines, we see that it is not an instance of (MT).

$$\begin{array}{l|ll} & p \Rightarrow \triangle_p q & \\ & \neg\triangle_\top q & \\ \# & \neg p & ?? \end{array}$$

A similar treatment could be developed for (2): once we represent the relevant operator's domain restriction, we see that the inference is not an instance of (∨E):

$$\begin{array}{l|ll} & p \vee \neg p & \\ & p \Rightarrow \triangle_p q & \\ & \neg p \Rightarrow \triangle_{\neg p} q & \\ \# & \triangle_\top q & ?? \end{array}$$

Similarly, perhaps, (1) can be understood as invalid by representing the reasoning as follows.[13] (RCM is a generally accepted principle of conditional logic, according to which the conditional is right upward monotone.)

$$\begin{array}{l|ll} & p \Rightarrow \triangle_p q & \\ & p \Rightarrow \triangle_p q \vee \triangle_{\neg p} q & \text{RCM} \\ & \neg p \Rightarrow \triangle_{\neg p} q & \\ & \neg p \Rightarrow \triangle_p q \vee \triangle_{\neg p} q & \text{RCM} \\ & (p \vee \neg p) \Rightarrow \triangle_p q \vee \triangle_{\neg p} q & \text{CA} \\ & \triangle_p q \vee \triangle_{\neg p} q & \text{MP} \\ \# & \triangle q & ?? \end{array}$$

An initial concern is that this rendering appears to misrepresent (1): a conditional of the form $(p \vee \neg p) \Rightarrow \triangle_p q \vee \triangle_{\neg p} q$ is not what is gotten by application of (CA) in (1). The relevant reading of (1c), rather, appears to be given by:

$$(p \vee \neg p) \Rightarrow \triangle_{p \vee \neg p} q$$

If that is right, then in order to understand the derivation of (1c) as an application of (CA) (as it in fact seems to be) the argument should be rendered as follows:

[13] An alternative representation might be offered, in which it is noted that $(p \vee \neg p) \Rightarrow \triangle_{p \vee \neg p} q$ does not follow from $p \Rightarrow \triangle_p q$ and $\neg p \Rightarrow \triangle_{\neg p} p$. This is akin to the strategy considered in §2.2.

$$
\begin{array}{l|ll}
 & p \Rightarrow \triangle_{p\vee\neg p}q & \\
 & \neg p \Rightarrow \triangle_{p\vee\neg p}q & \\
 & (p \vee \neg p) \Rightarrow \triangle_{p\vee\neg p}q & \text{CA} \\
 & (p \vee \neg p) & \top \\
 & \triangle_{p\vee\neg p}q & \text{MP} \\
\# & \triangle q & ??
\end{array}
$$

My concern[14] with *this* rendering is that, if modus ponens is indeed valid for indicative conditionals in natural language, we would expect $\triangle q$ to follow from $\triangle_{p\vee\neg p}q$: in other words, commitment to modus ponens for the natural language indicative conditional would seem to bring in tow commitment to the principle that $\triangle q$ follows validly from $\triangle_{p\vee\neg p}q$.

Note first that, in the context that I have described, there is the strong sense that the restricted operators—(6) and (7)—are simply *true*.

(6) It is better to smoke, assuming/given CANCER. $\triangle_p q$

(7) It is better to smoke, assuming/given ¬CANCER. $\triangle_{\neg p} q$

That's unsurprising, as (6) and (7) have *precisely* the logical forms that Kratzer (1981, 1991) that proposes for the conditionals (1a) and (1b): the compositional semantic function of the 'if'-clause is, on Kratzer's widely accepted account, *exhausted* by its restriction of some downstream operator.

This in mind, consider this schematic derivation:

$$
\begin{array}{|l}
\triangle_p q \\
\triangle_{\neg p} q \\
\triangle_{p\vee\neg p} q
\end{array}
$$

Given Kratzer's analysis, we may read derivations of this schematic form as representations of the logical form of many instances of (CA) in natural language. The relevant instance of this derivation appears to preserve acceptability in the context for (1): in this context, the acceptability of the following restricted operator seems to follow from the acceptability of (6) and (7).

(8) It is better to smoke, given CANCER or ¬CANCER. $\triangle_{p\vee\neg p} q$

This is as we would expect, given Kratzer's analysis, since (8) just is the logical form that Kratzer assigns (by default) to (1c).

The difficulty is that, if restricted operators are conditional in interpretation—as they surely *are*—and (MP) is valid for sentences of natural language with a conditional interpretation, we can derive the problematic conclusion, i.e., that it is better to smoke, from (8), given that CANCER ∨ ¬CANCER is a tautology.

Clearly (MP) does seem to govern the *vast majority* of cases of restrictable operators in natural language, and so there is reason to think that modus ponens should be formulated as a proof-theoretic principle for such operators (including the conditional operator $\Rightarrow$), along these lines:

$$
\begin{array}{|ll}
\triangle_\phi \psi & \\
\phi & \\
\triangle \psi & \Rightarrow\text{EO}
\end{array}
$$

[14]Another: there is no reason to suppose the logical forms of (1a) and (1b) are given by $p \Rightarrow \triangle_{p\vee\neg p}q$ and $\neg p \Rightarrow \triangle_{p\vee\neg p}q$. These logical forms are, it is true, the ones required to make the argument thus-represented go. But that is not a principled reason for thinking that the conditionals in question have these logical forms.

Here is just one example of an intuitively sound application of (MP) to a restrictable operator (it is obviously trivial to replicate).

(9) a. The likelihood that the marble is red, given that it is big, is 75%.
b. The marble is big.
c. ✓ So the likelihood that the marble is red is 75%.

This is as we would expect, if (i) (MP) is valid for natural language indicative conditionals, (ii) Kratzer is correct in thinking that (9a) is paraphrasable with the indicative conditional 'if the marble is big, there's a 75% chance it's red.'

One might worry whether there is a notion of logical consequence according to which (9) is valid, given that the truth of $\triangle_p q$ and p at w does not generally guarantee the truth of $\triangle q$ at w (on this point, see e.g. Charlow 2013b,c; Schulz 2018). Worry not: such an understanding has been independently suggested as a way of rescuing modus ponens from the challenge developed in Kolodny & MacFarlane (2010) (see e.g. Willer 2012; Yalcin 2012b; Bledin 2014). On this understanding, logical consequence is *informational* or *epistemic* in nature.[15] Valid arguments—what Bledin (2014) dubs "good deductive inferences"—are roughly, on this understanding, *knowledge*- or *information-preserving*.

Good deductive inference does not generally track preservation of truth at a world of evaluation, as a case like (9) shows. The fact that (i) a credal measure Pr that evaluates q as .75 likely given p at w, together with the fact that (ii) p is true at w does not imply that (iii) Pr evaluates q as .75 likely at w. However, (iii) does follow from (i), on the assumption that (ii') conditionalizing Pr on the information that p is idle, exactly as we would expect, if validity tracks the preservation of information, rather than the preservation of truth. More generally, although instances of (⇒EO) in natural language do not generally preserve truth with respect to a world of evaluation, they are ordinarily judged impeccable. Bledin's understanding of logical consequence would help to explain why: updating on the information expressed by the premises ordinarily makes available the information expressed by the conclusion.

Our difficulty is that, on this understanding, we would still apparently predict that a claim like (8) will allow us to infer the preferability of smoking by (⇒EO). Evidently, though, this is a *bad* deductive inference to draw. The information made available by (8) does not in any sense make available the information that smoking is preferred to not smoking. This appears to be precisely the problem with which we began, simply transposed into a metalanguage of restrictable conditional operators.

2.2 Context-Sensitivity Via Context-Shifting

A related strategy, due to Gillies (2009, 2010), assigns the claims familiar logical forms, but gives the conditional a distinctive *information-sensitive* and *context-shifting* interpretation.[16] On Gillies' analysis, in contrast with the context-sensitive analysis sketched in the prior section, (CA) is *not* generally valid.

I will motivate Gillies' analysis as he does, by considering epistemic modals (the epistemic necessity modal will be abbreviated '□', the epistemic possibility modal as '◇'). Consider a context where we know there is exactly one red or yellow marble concealed in an opaque box (Gillies 2010: 13). Each sentence in (10) is acceptable in this context.

(10) a. RED might be in the box and YELLOW might be in the box.
b. If it's not YELLOW, it must be RED.
c. If it's not RED, it must be YELLOW.

But Gillies proves that the sentences in (10) must be (contrary to fact) *inconsistent*, if we assign them their obvious logical forms (i.e., treat the epistemic necessity modals in (10b) and (10c) as taking narrow scope in their respective conditionals), assume Weak Centering for the conditional, and apply an otherwise standard semantics for conditionals and epistemic modals. Here are the relevant formal details:

[15] A theme of this paper, to which I will return toward the end, there is a difference in conceiving of logical consequence in *informational* terms and conceiving of it in *epistemic* terms (but that both conceptions are available to do theoretical work).

[16] I will not discuss dynamic accounts separately, since they, like Gillies' account, tend to avoid problematic predictions surrounding certain (∨E) arguments by invalidating (CA) (see, e.g., Willer 2012). I have already argued that invalidating (CA) does not help with (1) (and this section will make that argument again). While I think there is value in seeing how a theory like Gillies' can validate (MT), the uninterested reader should skip ahead to the next section.

Definition 1. *An information state i. is a function* $i : W \mapsto \wp(W)$ *(where* $w \in i_w$*)*

Definition 2. *A selection function* $f.(\cdot)$ *is a function* $f : W \mapsto (\wp(W) \mapsto \wp(W))$*, where* $f_w(p) \subseteq p$*.*

Definition 3. *A state* σ *is a pair* $\langle i, f \rangle$*, with i an information state, f a selection function, with f weakly centered:* $w \in p \rightarrow w \in f_w(p)$

Definition 4. *Let* $\sigma = \langle i, f \rangle$*, and let* $[\![p]\!]^{\sigma} = \{w : [\![p]\!]^{\sigma,w} = 1\}$*. Then:*

i. $[\![p \Rightarrow q]\!]^{\sigma,w} = 1$ *iff* $f_w(i_w \cap [\![p]\!]^{\sigma}) \subseteq [\![q]\!]^{\sigma}$

ii. $[\![\Box_p q]\!]^{\sigma,w} = 1$ *iff* $i_w \cap [\![p]\!]^{\sigma} \subseteq [\![q]\!]^{\sigma}$

iii. $[\![\Diamond_p q]\!]^{\sigma,w} = 1$ *iff* $i_w \cap [\![p]\!]^{\sigma} \cap [\![q]\!]^{\sigma} \neq \emptyset$

The inconsistency proof is straightforward (Gillies 2010: 18). Here is a sketch. Suppose the sentences of (10) are true at $\sigma = \langle i, f \rangle$ and w. There are two possibilities: w is a ¬YELLOW-world, or w is a ¬RED-world. If w is a ¬YELLOW-world, then by weak centering w is a □RED-world, hence every world compatible with i_w is a RED-world. If w is a ¬RED-world, then by weak centering w is a □YELLOW-world, hence every world compatible with i_w is a YELLOW-world. So either no YELLOW-worlds or no RED-worlds are compatible with i_w. Either way, (10a) is false at σ and w.

Gillies' proposed solution is elegantly simple: to revise the semantics for the conditional in Definition 4(i). For Gillies, an indicative antecedent has a dual compositional function: it introduces a (universal) quantifier over worlds selected by applying the selection to a range of possibilities compatible with the antecedent, while also *shifting the context of interpretation* for context-sensitive material in the consequent (to a context that incorporates the information expressed by the antecedent).

Definition 5. *The update of* $\sigma = \langle i, f \rangle$ *with p, notation* $\sigma[p]$*, is* $\langle \lambda w.i_w \cap [\![p]\!]^{\sigma}, f \rangle$

Definition 6. *Let* $\sigma = \langle i, f \rangle$*. Then* $[\![p \Rightarrow q]\!]^{\sigma,w} = 1$ *iff* $f_w(i_w \cap [\![p]\!]^{\sigma}) \subseteq [\![p]\!]^{\sigma[p]}$

Here are some attractions of this analysis that are worth noting. First, it avoids predicting the sentences in (10) inconsistent: if ¬YELLOW is true at σ and w, then, by weak centering, it follows that □RED is true at σ[¬YELLOW] and w, but does not generally follow that □RED is true at σ and w. We may not, then, draw the conclusion that every world compatible with i_w is a RED-world. This is sufficient to disrupt the proof of inconsistency sketched above.

Second, like the context-sensitivity strategy, it has the resources to blunt the force of Yalcin's counterexample to (MT) (though I doubt Gillies would endorse the strategy I am about to describe). Gillies' semantics naturally suggests the following understanding of logical consequence: logical consequence is preservation of truth *in the absence of illicit context-shift*. 'Illicit' context shift occurs, roughly, when, in the course of an argument, a *single sentence* is evaluated as false relative to σ and w, and true relative to σ' ($\neq \sigma$) and w. According to Gillies' semantics, in Yalcin's counterexample to (MT), a sentence of the form $\triangle q$ is evaluated as true relative to an *enriched* information state—an information state enriched with the information expressed by the antecedent of (5a). Meanwhile $\triangle q$ is evaluated as false relative to an unenriched information state—the global information state appropriate to the distribution of marbles Yalcin describes. There is *no antecedent reason* to formulate the notion of logical consequence in a way such that inference-patterns superficially in the mold of (MT) that *depend on this sort of context-shift* would be counted as valid (a point also well-made in Stojnić 2017).

On this strategy, instances of (MT) should be individuated *semantically* rather than *syntactically*, so that only inferences whose semantic contents instantiate the following set-theoretically valid "inference schema"—that is to say, inferences that *keep proper track of the context against which q is evaluated for truth*—count as bona fide instances of (MT):

$$[\![p \Rightarrow q]\!]^{\sigma} \qquad = \lambda w.f_w(i_w \cap [\![p]\!]^{\sigma}) \subseteq [\![q]\!]^{\sigma[p]}$$
$$[\![\neg q]\!]^{\sigma[p]}$$
$$[\![\neg p]\!]^{\sigma}$$

Proof. Let w satisfy $\lambda w.f_w(i_w \cap [\![p]\!]^\sigma) \subseteq [\![q]\!]^{\sigma[p]}$ and $[\![\neg q]\!]^{\sigma[p]}$. Hence, $w \notin [\![q]\!]^{\sigma[p]}$. Hence, $w \notin f_w(i_w \cap [\![p]\!]^\sigma)$. Given weak centering, $w \notin i_w \cap [\![p]\!]^\sigma$, hence (since $w \in i_w$) $w \notin [\![p]\!]^\sigma$, hence $w \in [\![\neg p]\!]^\sigma$. □

Inferences whose semantic contents instantiate this set-theoretic schema are valid instances of (MT). On Gillies' account, however, Yalcin's counterexample to (MT) instantiates the following *degraded* (because set-theoretically invalid) schema:

$$[\![p \Rightarrow q]\!]^\sigma \qquad = \lambda w.f_w(i_w \cap [\![p]\!]^\sigma) \subseteq [\![q]\!]^{\sigma[p]}$$
$$[\![\neg q]\!]^\sigma$$
$$\# \ [\![\neg p]\!]^\sigma$$

In my view this is a principled and compelling explanation of why Yalcin's (5) is invalid (and not, after all, appropriately regarded as an instance of MT). But now let us return to case (1). Is this explanation adaptable to it? Unlike the context-sensitive strategy, the context-shifting strategy easily yields a faithful representation of (1) (since the context-shifting strategy does not achieve its predictions by meddling with logical form). Notice further that (CA) fails on the context-shifting strategy. Suppose (10b) and (10c) are true. Then:

(11) a. $f_w(i_w \cap [\![\neg \textsc{yellow}]\!]^\sigma) \subseteq [\![\Box_\top \textsc{red}]\!]^{\sigma[\neg \textsc{yellow}]}$
b. $f_w(i_w \cap [\![\neg \textsc{red}]\!]^\sigma) \subseteq [\![\Box_\top \textsc{yellow}]\!]^{\sigma[\neg \textsc{red}]}$ Hence:

(12) a. $f_w(i_w \cap [\![\neg \textsc{yellow}]\!]^\sigma) \subseteq [\![\Box_\top \textsc{red}]\!]^{\sigma[\neg \textsc{yellow}]} \cup [\![\Box_\top \textsc{yellow}]\!]^{\sigma[\neg \textsc{red}]}$
b. $f_w(i_w \cap [\![\neg \textsc{red}]\!]^\sigma) \subseteq [\![\Box_\top \textsc{red}]\!]^{\sigma[\neg \textsc{yellow}]} \cup [\![\Box_\top \textsc{yellow}]\!]^{\sigma[\neg \textsc{red}]}$

It does *not* follow from this that the following conditions must hold. (Note that $i_w \cap [\![\neg \textsc{yellow} \vee \neg \textsc{red}]\!] = i_w$, hence that $\sigma[\neg \textsc{yellow} \vee \neg \textsc{red}] = \sigma$.)

(13) a. $f_w(i_w \cap [\![\neg \textsc{yellow} \vee \neg \textsc{red}]\!]) \subseteq [\![\Box_\top \textsc{red}]\!]^\sigma \cup [\![\Box_\top \textsc{yellow}]\!]^\sigma$
b. $[\![(\neg \textsc{yellow} \vee \neg \textsc{red}) \Rightarrow (\Box_\top \textsc{red} \vee \Box_\top \textsc{yellow})]\!]^{\sigma,w} = 1$

Something similar might be said to explain the badness of inference (1) (though I will not go through the details here): it relies on a specious application of (CA).

Nevertheless, although Gillies' semantics offers an elegant analysis of Yalcin's counterexample, our central concern remains: invalidating (CA) does not defeat the appeal of (1c). Where (1) goes wrong is in the move from (1c) to (1d); it is not in the move from (1a) and (1b) to (1c)! (I once again stress that I am not claiming that all applications of (CA) are good, only that the move from (1a) and (1b) to (1c) preserves truth or acceptability in its context of use.) The question of (CA)'s validity is beside the point. We presented (1c) as being "derived" from (1a) and (1b) by application of (CA), but this was not an essential feature of the presentation. (1c) is a directly acceptable way of *summarizing* the information contained in the relevant decision table; its truth is established *directly* by the relevant context, not via inference.[17] The truth of (1c), together with $p \vee \neg p$, does not warrant inferring (1d).

All this no doubt amounts to a puzzling state of affairs. It nevertheless seems to me the only way of reading of the data.

3 Taking the Case Seriously

Ultimately I will be arguing that such "failures" of (MP) are to be explained with a *normative*, rather than properly logical, account (on which the unsoundness of (1) is due to the fact that the way of representing the decision made salient in this context is an unreasonable way of representing a decision).

It would be easier if we could get by without a fancy normative story, while also taking this case seriously (by which I mean we take the judgment that (1c) is acceptable in its context of use as data, rather than something to be explained away). So that is where we will put our attention first. The best prospects here lie with a family of related views known as Alternative or Inquisitive Semantics (see, e.g., Alonso-Ovalle 2006; Groenendijk & Roelofsen 2009; Ciardelli & Roelofsen 2011). This section provides a general overview of such views, though it ultimately zeroes in the theory of Bledin (2020). Although

[17] Note that Gillies' semantics renders it false, if the basic context (state) of evaluation provides a preference ordering according to which not smoking is better than smoking.

Bledin's account does better with respect to (1) than any other account we have so far considered, I will suggest that his account *under-generates* good deductive inferences (in the broader setting of what I will call a logic of decision). The theory I go on to state will attempt to find a middle ground on which certain intuitively sound episodes of dominance reasoning are deductively licensed, although not of course the episode of dominance reasoning dramatized in (1).

3.1 Inquisitive Logic

Let us begin with a very basic observation: (the relevant reading of) $(p \vee \neg p) \Rightarrow q$ plausibly entails both $p \Rightarrow q$ and $\neg p \Rightarrow q$. Combine this with (CA), and we have:

$$(p \vee \neg p) \Rightarrow q \dashv\vdash (p \Rightarrow q) \wedge (\neg p \Rightarrow q)$$

If $(\vee E)$ is not (as many believe) a valid rule of inference, we cannot derive q from the premise set Γ_1. Unsurprisingly, q is not generally derivable from a *logically equivalent* premise set Γ_2.

$$\Gamma_1 = \{(p \Rightarrow q) \wedge (\neg p \Rightarrow q), p \vee \neg p\}$$
$$\Gamma_2 = \{(p \vee \neg p) \Rightarrow q, p \vee \neg p\}$$

There is no doubt something to this. Strikingly, however, it *appears* to concede that indicative conditionals of the form $(p \vee \neg p) \Rightarrow q$ do not generally go in for (MP) (while supplying an explanation of this fact).[18] This prompts an obvious question: under exactly what conditions *is* inferring q from $(p \vee \neg p) \Rightarrow q$ logically permitted?

Alternative/Inquisitive Semanticists may see an answer ready to hand. Following this tradition, let us distinguish between simplifying and non-simplifying representations of a conditional of surface form $(p \vee \neg p) \Rightarrow q$. On the simplifying representation, the antecedent of a conditional $(p \vee \neg p) \Rightarrow q$ denotes an inquisitive (question-like) content, namely the set of propositional alternatives $\{p, \neg p\}$. On the non-simplifying representation, the antecedent denotes a proposition, derived by application of a flattening operator ! (such that, if S is a set of alternative propositions, $!S = \bigcup S$) to $\{p, \neg p\}$.

	$\{p, \neg p\} \Rightarrow \triangle q$	CA
	$!\{p, \neg p\}$	$\top$
#	$\triangle q$	??

	$\{p, \neg p\} \Rightarrow \triangle q$	CA
	$\{p, \neg p\}$	??
#	$\triangle q$	MP

The template on the right does seem to be an available representation for the cases under consideration. A speaker can indicate that the intended reading of the bare disjunction denotes a set of alternatives $\{p, \neg p\}$, rather than $!\{p, \neg p\}$, simply by, e.g., choosing a different pattern of focal stress.

(14) A: Are you a **student or teacher**? $?!\{s, t\}$
B: Yes/No

(15) A: Are you a **student** or **teacher**? $?\{s, t\}$
B: I'm a student/I'm a teacher [#Yes/#No]

The latter pattern of focal stress seems optional (perhaps preferred) in, e.g., (1).

(16) You **will get cancer or won't**.

(17) You'll either **get cancer** or **won't**.

If the template on the right is an available representation, it is essential to this line of reply that there is no rule of deductive inference that allows one to introduce a sentence expressing $\{p, \neg p\}$ at this step of the argument. Here is a natural (if slightly naïve) question: what could justify this sort of prohibition? In introducing $\{p, \neg p\}$, one is introducing something *informationally trivial*, but with inquisitive or alternative-presenting content (Groenendijk & Roelofsen 2009). It is natural to regard such a move as

[18] Bledin (2020) would resist this characterization of the theory, for reasons we will see below.

deductively licensed: in introducing $\{p, \neg p\}$, one is, roughly, deciding to consider or raise to salience a question (is p true, or $\neg p$?). But, first, this is a question that is surely *already salient* in the contexts I have described, and so it cannot be a distortion to represent the set of alternatives $\{p, \neg p\}$ explicitly. Further, it seems that it is always possible to introduce a set S that partitions the relevant possibilities *without collateral epistemic work*, since, when S partitions the relevant possibilities, introducing S will *introduce no new information* into the discourse. And so the introduction of $\{p, \neg p\}$ would seem to be deductively permitted, even if such a question were not already salient in the relevant contexts.[19]

Bledin (2020), building on Ciardelli (2018), uses an Inquisitve Semantic framework to state an account of cases like (1) that answers this naïve challenge. Bledin's core suggestion is this: if we represent an agent as reasoning from an inquisitive premise (or supposition) $?p$, we thereby represent them as having an (*indeterminate*) opinion on the question $?p$.

> *How much is one Bitcoin worth in US dollars?* stands in for one of the various ways in which the question expressed by this sentence can be resolved. When this constituent interrogative appears as a premise in an inference, we can come to learn things about any information state that settles it... For example, if we infer the polar interrogative *Is one Bitcoin worth more than a thousand dollars?*, we thereby establish that the more specific type of information that settles how much one Bitcoin is worth in US dollars yields the less specific type of information that settles whether one Bitcoin is worth more than a thousand dollars.
>
> ...[W]e can think of interrogatives as placeholders for arbitrary decision states of a given type. In [(1)], the premise ['I will or will not get cancer'] stands in for an arbitrary decision state that settles the question of whether you are going to [get cancer]. (149)

On the Ciardelli-Bledin account, the argument from $?p$ and $?p \Rightarrow \triangle q$ to $\triangle q$ is to be regarded as valid, in this rough sense: a decision state that satisfies the condition on decision states expressed by $?p$ and $?p \Rightarrow \triangle q$ *is* a decision state that accepts $\triangle q$. An agent who merely entertains the question *is it the case that p?* is not, in general, to be understood as satisfying the condition on decision states expressed by $?p$: satisfying this condition, in the logically relevant sense, implies either accepting p or accepting $\neg p$. On this account, for an agent to accept (satisfy the condition on decision states) $?p \Rightarrow \triangle q$ is for them to be such that both:

- If they accept p, they must accept $\triangle q$.
- If they accept $\neg p$, they must accept $\triangle q$.

It is obvious that, on this account, accepting $?p \Rightarrow \triangle q$ does not require accepting $\triangle q$, since an agent can accept $?p \Rightarrow \triangle q$ without thereby accepting p or accepting $\neg p$.[20]

Bledin's account aims to answer both of the following questions:

- Why does the argument in (1) fail?
- Under what conditions is inferring $\triangle q$ from $(p \vee \neg p) \Rightarrow \triangle q$ logically permitted? What accounts for the contrast between good and bad inferences of this form?

According to Bledin, what distinguishes good instances of this form, like (18), from bad, like (1), is a

[19] This isn't to say one can go around freely expressing such contents in discourse. Questions are typically *not relevant* unless they address a larger question within an extant discourse (see, e.g., Roberts 1996), and speakers must generally make their contributions relevant, vis-à-vis the broader goal of the extant discourse. This would appear to be orthogonal to the logical issues under consideration here.

[20] Ciardelli and Bledin do not say much by way of explicating the idea of a theory of logically valid reasoning making use of *premises* whose content is *indeterminate*. I tend to understand their account of logical consequence supervaluationally: ψ is a logical consequence of ϕ just when, for any *determinate assignment* g of content g such that $g(\chi)$ is a semantic alternative for χ, $g(\psi)$ is a logical consequence of $g(\phi)$. (Ciardelli compares his account to the Intuitionistic treatment of disjunction, which is also an appropriate comparison.) Even with this understanding, the account has puzzling features: for the reasons canvassed above, one would naïvely think it should be possible to represent an agent as reasoning validly from a disjunction $p \vee \neg p$, interpreted inquisitively as $\{p, \neg p\}$, without thereby representing that agent as taking a stance on whether p. The account I will ultimately sketch does not share this puzzling feature.

distinctive semantic property of the consequent.

(18) a. If it's raining, the match is likely be cancelled.
b. If it's not raining, the match is likely be cancelled.
c. So, if it's raining or it isn't, the match is likely be cancelled.
d. So, the match is likely be cancelled.

Specifically, Bledin says an inference of this form is good just when the consequent $\triangle q$ has a property he dubs Coarse Distributivity. (Note: here $s \subseteq W$, while $\preceq$ is a weak partial order on subsets of s.[21])

Coarse Distributivity (Bledin 2020: 147)

ϕ is Coarsely Distributive iff for any decision state $\langle s, \preceq\rangle$ and partition $\{s_1, ..., s_n\}$ of s, if $[\![\phi]\!]^{s_i,\preceq} = 1$ (for all $1 \leq i \leq n$), then $[\![\phi]\!]^{s,\preceq} = 1$.

Informally, ϕ is Coarsely Distributive just when acceptance of ϕ relative to every possibility in a partition implies outright (unconditional) acceptance of ϕ. For Coarsely Distributive ϕ, Bledin's account predicts it is impossible for an agent to accept $?p \Rightarrow \phi$ without also accepting ϕ.

Bledin's thought here is that sentences expressing comparative preference are a paradigm of Non-Distributivity: conditional on the proposition that I get cancer (from smoking), as well as conditional on the proposition that I do not get cancer (from smoking), it is better for me to smoke; nevertheless, unconditionally, it is obviously not the case that it is better for me to smoke. In stark contrast, sentences expressing comparative probability ('ϕ is likelier than ψ') are a paradigm of Coarse Distributivity. Note first the following fact about probability/comparative likelihood.

Likelihood is Coarsely Distributive

Consider any state $\langle s, \preceq\rangle$ and partition $\{s_1, ..., s_n\}$ of s. If, for each i, $q \cap s_i$ is likely in $\langle s, \preceq\rangle$ ($:= \neg q \cap s_i \prec q \cap s_i$), then q is likely in $\langle s, \preceq\rangle$ ($:= \neg q \prec q$).

This follows from a more general fact about probability measures:

Statewise Probabilistic Dominance (SPD)

Consider any s and $\{s_1, ..., s_n\}$ that partitions s, and let Pr be a probability measure. If $\Pr(q \mid s_i) \geq x$ (for all $1 \leq i \leq n$), then $\Pr(q \mid s) \geq x$.

Informally, SPD says that, if q is likely conditional on each cell of a partition of s, then q is likely conditional on s.

For Bledin, there is no strictly *logical* justification of inference (18).[22] It is *not* an instance of modus ponens (since for it to be so, we would need to appeal to {it's raining, it's not raining} as a premise, which, for reasons already seen, Bledin's notion of logical consequence will prohibit). To see this more clearly, it is helpful to examine a case where a formally indistinguishable inference fails. A marble has been drawn from an urn containing 100 marbles in the following distribution:

- 30 completely red marbles (all of these small)
- 30 completely blue marbles (all of these small)
- 40 half-red and half-blue marbles (all of these big)

In this case, the alternatives in {it's red, it's blue} ('it's red' here understood to mean that it's somewhere red, 'it's blue' that it's somewhere blue) are exhaustive: all the marbles are somewhere red or blue.

(19) a. If it is red, it is likely big.
b. If it is blue, it is likely big.
c. So, if it is red or it is blue, it is likely big.
d. #So, it is likely big.

[21] As the ensuing discussion illustrates, an ordering like $\preceq$ may be used to represent a binary preference relation, as when we are interpreting a construction of comparative preference ('better'), but also to represent a binary probabilistic relation, as when we are interpreting a construction of comparative probability ('likely' or 'likelier than not').

[22] Another way of putting this point is that Bledin is not interested in providing a *proof theory* according to which (18) is a deductively permitted inference. More on this below.

Although (18) and (19) share a logical form, (18) is impeccable and (19) is outrageous. The only difference between (18) and (19) is of a substantive character: the set of alternatives in (18) is a *partition* of the set of relevant possibilities, while the set of alternatives in case (19) is not (since the elements of {it's red, it's blue}, though jointly exhaustive, are also jointly compatible). Although neither (18) nor (19) amounts to an instance of modus ponens, (18) is nevertheless valid, its validity owing to a substantive (necessary) truth about comparative likelihood (rather than to modus ponens).

3.2 *Reasoning by Dominance*

Bledin's inquisitive logic is elegant and explanatory, but I will suggest that it is ill-suited as a logic of dominance reasoning. Here I will be making the case that (1) fails *qua* inference because it is Spurious Reasoning by Dominance—not because constructions of comparative preference ('smoking is preferred to not smoking') generally fail to be Coarsely Distributive. The contrast between Spurious and Non-Spurious Reasoning by Dominance is ultimately, I will suggest, to be accounted for in the same way as the contrast between (19) and (18): logical rules, like modus ponens, may be understood as rules of *rational inference*, and the theory of rational inference may build in *pragmatically* (as opposed to logically or grammatically) justified assumptions about rational ways of representing one's information.

As decision theorist readers no doubt noticed right away, (1) is a textbook case of *spurious* reasoning by Dominance. Reasoning by Dominance can be initially (but, we will see, not altogether satisfactorily) characterized with the following principle:

Statewise Dominance (SD)

Consider decision problem Π. (s_i is a relevant contingency, a_j an available action, $\text{Val}_{\leq}(a_j \mid s_i)$ the degree to which performing a_j, given s_i, satisfies the preferences encoded in $\leq$, which I will abbreviate as its 'value'.)

Π	s_1	...	s_n
a_1	$\text{Val}_{\leq}(a_1 \mid s_1)$	...	$\text{Val}_{\leq}(a_1 \mid s_n)$
...	...	...	...
a_m	$\text{Val}_{\leq}(a_m \mid s_1)$	...	$\text{Val}_{\leq}(a_m \mid s_n)$

If $\text{Val}_{\leq}(a_k \mid s_i) < \text{Val}_{\leq}(a_j \mid s_i)$ (for all $k \neq j$ and $1 \leq i \leq n$), then $\text{Val}_{\leq}(a_k) < \text{Val}_{\leq}(a_j)$ (for all $k \neq j$).

The idea behind the SD principle is simple: if, for every relevant contingency, the value of *a* given that contingency exceeds the value of any other available action given that contingency, then the unconditional value of *a* must exceed the unconditional value of any other available action. Nevertheless, cases of "spurious" reasoning by Dominance, like (1), show that SD cannot hold in full generality: if we allow Statewise Dominance to apply to the decision problem associated with (1), we fail to take into account the (causal or evidential) dependence[23] of relevant contingencies on the available actions. Restricting SD to cases in which the relevant contingencies are *independent* of the available actions is essential, if our goal is to rationalize any sort of action whose performance is, by some metric, costly to an agent, but which nevertheless *contributes to* their prospects for avoiding some undesirable outcome—e.g., national defense expenditures (Jeffrey 1983: 8–9), payment of protection money (Joyce 1999: 115ff), and, indeed, refraining from smoking.

In decision-theoretic contexts, this sort of restriction is often understood as a restriction of the dominance principle to "*well-formed*" decision-problems: decision problems in which relevant contingencies are, simply, independent of the available actions.[24]

[23]The question of how to understand dependence, in the sense relevant for formulating a precise representation of reasoning by Dominance leads headlong into the debate between Causal and Evidential formulations of Decision Theory. Probabilistic understandings of the problematic dependence are roughly associated with "Evidential" Decision Theories, causal understandings with "Causal" Decision Theories (see esp. Gibbard & Harper 1981; Joyce 1999). More on this below.

[24]The stipulation that a decision theory's choice function is defined only over well-formed decision problems is most closely associated with Savage (1972). Alternatives to Savage's theory, like Jeffrey (1983), also restrict dominance reasoning to a certain class of decision problems: those in which the relevant states are states (that the agent regards as) evidentially independent of the

Well-Formedness

Π is well-formed only if for each i, j: s_i is independent of a_j.

This notion to hand, a restricted version of SD can be stated as follows:

Restricted Statewise Dominance (RSD)

For *well-formed* Π: if $\text{Val}_{\leq}(a_k \mid s_i) < \text{Val}_{\leq}(a_j \mid s_i)$ (for all $k \neq j$ and $1 \leq i \leq n$), then $\text{Val}_{\leq}(a_k) < \text{Val}_{\leq}(a_j)$ (for all $k \neq j$).

On the account I propose here, (1) fails roughly because the decision problem I made salient in describing the case is malformed, not because Coarse Distributivity fails for the comparative preferability operator.

Is it really the nature of the relevant decision situation that accounts for (1)? I'll say yes,[25] but let us look at a reason to think not. Consider this well-known case. The context is the Miners scenario from Kolodny & MacFarlane (2010): ten miners are trapped, in Shaft A or Shaft B, both of which are rapidly filling with water. Blocking the shaft the miners are in will save all ten; blocking the wrong shaft will kill all ten; blocking neither will save nine, killing one.

(20)
a. If the miners are in shaft A, blocking a shaft is best.
b. If the miners are in shaft B, blocking a shaft is best.
c. So, if the miners are in shaft A or in shaft B, blocking a shaft is best.
d. ??So, blocking a shaft is best.

In this case, clearly, there is no failure of act-state independence to appeal to: the miners are where they are, independent of what we choose to do. (20) is not explained by appeal to RSD; what reason is there to think that (1) is?

Here, though, is a relevant difference between (1) and (20). In (20) the context *guarantees a true reading of the conclusion*: there is a clear sense in which either blocking A is (unconditionally) preferable or blocking B is (unconditionally) preferable, which of these depending on the miners' actual location. I'll henceforth refer to this sense of preferability as *primary preferability*.

There is also evidently a sense of 'best'—which we might roughly gloss as 'selected by the correct decision theory', and to which I will refer as *deliberative preferability*—according to which blocking a shaft is *clearly not* best. Reading 'best' this way probably accounts for our squeamishness with respect to (20d) (appearing, as it does, at the end of an implicit episode of practical deliberation). To screen off this sense in reading (20), let me reiterate an implicit supposition: *we have no idea where the miners are*. Evidently, (20a) still has a true reading, even though, even on the supposition that the miners are in A, no plausible decision theory will, on the supposition that we have no idea where the miners are, recommend blocking a shaft. I conclude that when we judge (20a)–(20c) acceptable in the miners context, we are reading 'best' as primary preferability.

Primary preferability does appear to coarsely distribute, if not in full generality, then certainly in contexts like (20). And this raises a question that Bledin's account does not directly answer: if primary preferability obeys coarse distributivity in contexts like (20), why does it fail to in contexts like (1)? Notice that in case (1), it would *not* be best if you smoked—or, at least, nothing about the context suffices to guarantee that *smoking is preferred to not smoking*, on *any* way of reading this claim. The context in (1) does not guarantee that the best possibilities (within your practical or causal reach) are possibilities in which you smoke: *smoking might give you cancer*, in which event it would have been better not to smoke. By contrast, the context in (20) does guarantee that the best possibilities (within your practical or causal reach) are possibilities where you block a shaft. This is a clear point of contrast between (1) and (20), explained precisely by the failure of act-state independence in (1).

---

actions (the agent regards as) available (to her). By "well-formed decision problem", I mean simply a decision problem in which dominance reasoning is well-applied. For an overview of the dialectic, see esp. Joyce (1999: Ch. 5).

[25] This is also the diagnosis that Gibbard & Harper (1981); Cantwell (2006) offer for the unsoundness of (∨E) dominance arguments like (2).

4 Logic of Decision

This section will begin by stating a toy model theory for a toy language $\mathcal{L}$ built from a base propositional language, binary likelihood operator ▲, binary preference operator ⋆, and two-place conditional operator ⇒.[26] I will use this toy theory to develop and illustrate a thesis about logical consequence, according to which *normative questions encroach on questions about how to formalize good deductive reasoning*. I will ultimately propose that there is a reasonable understanding of logical consequence, according to which logical consequence might be regarded as *relative to*, or *dependent upon*, a substantive normative (e.g., decision) theory.

4.1 A Toy Theory

I begin by refining the notion of a decision state.

Definition 7. *A **decision state*** $\Pi = \langle \mathcal{S}, \mathcal{A}, \mathrm{Pr}, \mathrm{Val}_{\leq} \rangle$ *where:*

- $\mathcal{S} = \{s_1, ..., s_n\}$ *is a set of states represented in* Π.
- $\mathcal{A} = \{a_1, ..., a_m\}$ *is a set of actions represented in* Π.
- $\mathrm{Pr}(\cdot \mid \cdot)$ *is a conditional credence function, where* $\mathrm{Pr}(\cdot) := \mathrm{Pr}(\cdot \mid \top)$.
- $\mathrm{Val}_{\leq}(\cdot \mid \cdot)$ *is a conditional value function, where* $\mathrm{Val}_{\leq}(a \mid s)$ *is the degree to which a is preferred conditional on s and* $\mathrm{Val}_{\leq}(\cdot) := \mathrm{Val}_{\leq}(\cdot \mid \top)$.

Definition 8. $\Pi[\phi]$*, the **update** of* $\Pi = \langle \mathcal{S}, \mathcal{A}, \mathrm{Pr}, \mathrm{Val}_{\leq} \rangle$ *on* $\phi \in \mathcal{L}$*, is the result of conditioning* $\mathcal{S}$*,* Pr*, and* Val *on x:* $\Pi_\phi = \langle \mathcal{S}[\phi], \mathcal{A}, \mathrm{Pr}(\cdot \mid \phi), \mathrm{Val}_{\leq}(\cdot \mid \phi) \rangle$.

$$\mathcal{S}[p] = \{s \in \mathcal{S} : s \subseteq p\}$$
$$\mathcal{S}[\neg\phi] = \mathcal{S} - \mathcal{S}[\phi]$$
$$\mathcal{S}[\phi \wedge \psi] = \mathcal{S}[\phi] \cap \mathcal{S}[\psi]$$
$$\mathcal{S}[\phi \vee \psi] = \mathcal{S}[\phi] \cup \mathcal{S}[\psi]$$

Like Bledin (as well as Cariani et al. 2013; Charlow 2016, 2018), we define satisfaction for $\mathcal{L}$ relative to a decision state $\Pi = \langle \mathcal{S}, \mathcal{A}, \mathrm{Pr}, \mathrm{Val}_{\leq} \rangle$. Satisfaction in general is *idle update*, while entailment is simply *preservation of satisfaction* with respect to an arbitrary decision state:

$$\Pi \vDash \phi \text{ iff } \Pi[\phi] = \Pi$$
$$\phi_1, ..., \phi_n \vDash \psi \text{ iff } \forall \Pi \vDash \phi_1 ... \vDash \phi_n : \Pi \vDash \psi$$

I won't define updates for ▲ or ⋆ or ⇒,[27] though I will provide their satisfaction conditions. Let $[\![\phi]\!]^\Pi$ designate the set of states represented in Π that support ϕ: $[\![\phi]\!]^\Pi = \{s \in \mathcal{S} : s \subseteq \phi\}$.[28]

$$\Pi \vDash \blacktriangle(\psi \mid \phi) \text{ iff } \mathrm{Pr}([\![\psi]\!]^\Pi \mid [\![\phi]\!]^\Pi) > \mathrm{Pr}([\![\neg\psi]\!]^\Pi \mid [\![\phi]\!]^\Pi)$$
$$\Pi \vDash \star(\psi \mid \phi) \text{ iff } \mathrm{Val}_{\leq}([\![\psi]\!]^\Pi \mid [\![\phi]\!]^\Pi) > \mathrm{Val}_{\leq}([\![\neg\psi]\!]^\Pi \mid [\![\phi]\!]^\Pi)$$
$$\Pi \vDash \phi \Rightarrow \psi \text{ iff } \forall s \in [\![\phi]\!]^\Pi : \Pi[s] \vDash \psi$$

$\mathrm{Pr}([\![\psi]\!]^\Pi \mid [\![\phi]\!]^\Pi)$ is naturally defined as the sum of the probabilities (given ϕ) for the possibilities in $\mathcal{S}$ that entail ψ: $\mathrm{Pr}([\![\psi]\!]^\Pi \mid [\![\phi]\!]^\Pi) := \mathrm{Pr}(\bigcup [\![\psi]\!]^\Pi \mid \bigcup [\![\phi]\!]^\Pi)$

Supposing ⋆ is primary preferability, $\mathrm{Val}_{\leq}([\![\psi]\!]^\Pi \mid s)$ represents, to a first pass, the degree of value an agent whose decision situation is representable with $\Pi[s]$ can intentionally cause or realize by making it

[26] In spite of the qualifications (toy semantics, toy language), this is not far from the sort of semantic theory I would defend for this fragment of English. One assumption I implicitly make for the sake of presentation is that these operators take only sentences from the base propositional language as restrictors.

[27] Were I to do so, I would simply treat them as tests, in the sense of Veltman (1996): updates that return the original state in the event that the sentence is satisfied, and which return an absurd state otherwise (compare Charlow 2015).

[28] I am here assuming that $[\![\phi]\!]^\Pi$ and $\Pi[\phi]$ are defined iff, for all $s \in \mathcal{S}$, $s \cap \phi = s$ or $s \cap \phi = \emptyset$ (that is to say, ϕ presupposes, in a loose sense, that the relevant information is partitioned ϕ-wise). This is a visibility presupposition, in the sense of Yalcin (2011).

the case that ψ (ψ-ing, for short). Primary preferability is here construed akin to (relevantly) informed advisability: if you are "lost in the woods without a map or compass," the primarily preferable thing—the thing a relevantly informed advisor would advise you to do—would be to walk in the direction of your car (whereas the deliberatively preferable thing to do is to "pursue one of the standard strategies for getting out of trackless woods: walk carefully in a straight line by sighting along trees, or go consistently downhill") (Gibbard 1990: 18–19).

To elucidate the notion, consider this representation of the decision situation relevant for (20):

	they're in A	they're in B
block A	10 live	10 die
block B	10 die	10 live
don't block A or B	9 live	9 live

A central fact of this decision situation is that, whatever the miners' location, the degree of value you're positioned to bring about if you block a shaft exceeds the degree of value you're positioned to bring about if you don't block a shaft. Blocking a shaft is what you would choose to do, if only you knew where the miners were—something that is *guaranteed by* the fact that blocking a shaft is what you would choose to do if you had settled which state of your decision situation was actual.

This case supports the conditional $(A \vee \neg A) \Rightarrow \star BL$, read: if the miners are in A or aren't in A, blocking a shaft is (primarily) preferable. On the assumption that dominance reasoning to conclusions about primary preferability is logical—on the assumption that $\star$ is coarsely distributive—we predict (it would appear correctly) that $\star BL$ is a logical consequence of $(A \vee \neg A) \Rightarrow \star BL$.

$$\begin{aligned}\Pi \vDash (A \vee \neg A) \Rightarrow \star BL \text{ implies } & \forall s \in [\![A \vee \neg A]\!]^{\Pi} : \Pi[s] \vDash \star BL \\ \text{implies } & \Pi[A] \vDash \star BL \text{ and } \Pi[\neg A] \vDash \star BL \\ \text{implies } & \Pi \vDash \star BL\end{aligned}$$

But, of course, on the assumption that $\star$ is coarsely distributive, the present theory also predicts (it would appear incorrectly) that (1d) is a logical consequence of (1c), and therefore that smoking's degree of primary preferability exceeds that of not smoking. I'll now consider two ways of modifying the basic theory to avoid this false prediction (while also managing to construe dominance reasoning as logical), each of which answers to a different set of theoretical desiderata for what I will call a "logic of decision."

4.2 Logic Bound

Consider first a very direct fix: we stipulate that the theory we have provided is a theory covering all, but *only*, the reasonable decision states (provided that a reasonable decision state is defined as a decision state witnessing act-state independence):

Definition 9. *A decision state* $\Pi = \langle \mathcal{S}, \mathcal{A}, \Pr, \mathrm{Val}_{\leq} \rangle$ *is* ***reasonable*** *only if, for all* $s \in \mathcal{S}$ *and* $a \in \mathcal{A}$, $\Pr(s \mid a) = \Pr(s)$.

The precise stipulation is that the semantic and logical clauses provided above are *defined for all and only reasonable* Π. The decision states that our semantic theory admits are reasonable, by convention or assumption.

Cases like (1) are theorized broadly as follows. The case makes salient a (type of) decision state that is unreasonable; but the theory is, strictly speaking, silent on which inferences are deductively licensed from such a decision state.

On this theory, dominance reasoning (to conclusions about primary preferability) is deductively licensed *depending* on (normatively significant) features of the practical context. Notice that—for reasons already seen—for any reasonable decision state Π such that $\Pi \vDash$ (1c), $\Pi \vDash$ (1d). Imagine (obviously contrary to fact) that one's risk of cancer was *independent* (or was *represented as* independent) of smoking. In such a practical context, the reasoning in (1) would of course be impeccable.

(21) a. If you get cancer or you don't, it's better to smoke.
b. ✓ So, it's better to smoke.

On the present account, the contrast is explained by a shift in the properties of the relevant decision problem: the decision state used for evaluation of (21) is, by assumption, reasonable, and so the conclusion must be satisfied by any state that satisfies the premise. For this reason, the argument in (1) may be regarded as a *Strawson-valid* instance of modus ponens, since for any *reasonable* decision state that satisfies (1c), that decision state satisfies (1d).[29]

With one further natural stipulation, this theory also yields an account of dominance reasoning in the more familiar sense (which I will suppose is represented as reasoning from (exhaustive) premises about primary preferability to conclusions about deliberative preferability, the latter of which I'll represent using the operator ☆).

Means-End Coherence (MEC)

For all Π, ϕ, a: if a is available in Π and $\Pi[\phi] \vDash a$: $\Pi \vDash \star\phi$ implies $\Pi \vDash ☆a$.

MEC says that if ϕ is (primarily) preferred and there is an available action a such that ϕ implies doing a, then a is (deliberatively) preferred. To illustrate informally:

- In the Miners case, blocking a shaft is primarily preferred, roughly because it's the only way to save all ten miners. But the available actions are: block A, block B, or block neither; no available action is such that blocking a shaft implies doing that action, and so MEC doesn't require that you deliberatively prefer to block a shaft.
- To contrast, in the modified smoking scenario (21), MEC does require that you deliberatively prefer smoking, since smoking is (i) primarily preferred and also (ii) identical to an action that is available to you.

What is the logical status of modus ponens—which I will (recalling §2.1) formulate as a "proof-theoretic" principle for all binary operators $\triangle$—in such a theory?

$$\phi, \triangle_\phi \psi \vdash \triangle\psi$$

I have here been trying to formulate a theory on which such a proof-theoretic principle could be seen as sound with respect to $\vDash$. The present theory does not, however, quite make good on this goal: recalling case (19), the following probabilistic inference is logically acceptable when ϕ and ψ are represented as incompatible, but not otherwise.

$$\phi \vee \psi, (\phi \vee \psi) \Rightarrow \blacktriangle\chi \vdash \blacktriangle\chi$$

One, again very direct, possibility for enforcing this is simply to expand the criteria of reasonability, along the following lines:

Definition 10. *A decision state* $\Pi = \langle \mathcal{S}, \mathcal{A}, \mathrm{Pr}, \mathrm{Val}_\leq \rangle$ *is **reasonable for*** $(\phi \vee \psi)$ *only if* $\mathcal{S}[\phi] \cap \mathcal{S}[\psi] = \emptyset$.

At this point it would be natural to ask what is supposed to justify the assumption that the decision states used to provide a logic (and model theory) of decision must be reasonable in both of these senses. The assumption is, I'll suggest, best regarded as a *normative idealization* (very much akin, as I will explain, to the assumption that Pr is, by definition, a *probability* measure).[30]

---

[29] Strawson validity is the property an argument has when for any state of evaluation σ, if σ satisfies the presuppositions of the argument's premises and satisfies the premises, σ satisfies the argument's conclusion (see von Fintel 1999). In calling such an inference Strawson-valid, I am here thinking of reasonability as a kind of presupposition (although not, of course, in the exact sense invoked in linguistic theory).

Since $\vDash$ does not relate unreasonable decision states to sentences, the theory does not say whether the decision state Π made salient in (1) is such that $\Pi \vDash$ (1c). So this theory does not, strictly speaking, account for the judgment that (1c) is true/acceptable/satisfied, when assessed from Π. Notice however that Π and (1c) do stand in a relation that we might call *Strawson-satisfaction*: for any modification Π' of Π that is reasonable and is otherwise like Π, $\Pi' \vDash$ (1c) and $\Pi' \vDash$ (1d). More on this shortly.

[30] It is significant, for the sake of comparing the theories, that Bledin (2020), following Hamblin (1958); Groenendijk & Stokhof (1984); Biezma & Rawlins (2012), assumes that a semantically well-formed question (here understood to include an inquisitively interpreted disjunction) is such that its answers *must be* understood as mutually exclusive. Although this assumption *entails* that a decision state Π can satisfy $\phi \vee \psi$ only if $\mathcal{S}[\phi] \cap \mathcal{S}[\psi] = \emptyset$—and thus has basically the same effect as supposing that abstract representations of states of decision are reasonable for semantically well-formed disjunctions, as a matter of definition—I nevertheless do not believe it is warranted, in light of cases like (19).

Philosophers who model the notion (and logic) of degreed belief commonly assume, as I have done here, that a state of degreed belief consists in an agent's relation to a probability measure (or set thereof) that is updated by via (pointwise) conditionalization.[31] But the assumption that probability measures are (i) probabilistically coherent and (ii) are updated via conditionalization is standardly justified by appeal to a *normative argument* (although the style of argument varies with the explanatory interests of the theorist): probabilistic incoherence or failure to conditionalize "subjects" or "disposes" an agent to epistemically undesirable inferences (in the context of Accuracy-Dominance arguments for probabilistic coherence or conditionalization) or practically undesirable inferences (in the context of Dutch Book arguments for probabilistic coherence or conditionalization).[32]

Here, too, the requirement that a decision state witness act-state independence (as well as the assumption that a decision state treat disjunctions as exclusive) can be seen as motivated by broadly normative considerations: failures of reasonability subject or dispose an agent to undesirable inferences, and, for certain modeling purposes, it makes sense to idealize away such failures (as when we idealize away incoherent credences when modeling certain features of probabilistic talk and thought). A theory that assumes Reasonability and Means-End Coherence yields a model theory and theory of deductive inference for a certain type of *minimally rational agent*—roughly speaking, an agent whose decision situation is representable as reasonable, and who takes means appropriate to their ends. If you are in need of a logic of practical talk and thought, on which canonical forms of deductive practical reasoning (e.g., dominance reasoning) can represented *as logical*, here is a theory you can put to real use.

4.3 Logic Unbound

A modeling strategy of this nature does raise something of a puzzle: in what sense can an incoherent or unreasonable agent be said to be "subject" to an inference, if the logic and model theory are defined only over representations of coherent or reasonable agents? I take the thrust of this question to be that we have need of a notion of *logical commitment*, according to which an agent who represents their situation in an unreasonable or incoherent way can be represented as logically committed to an unreasonable conclusion.

Dutch Book arguments help to illustrate what I have in mind. As recent commentaries (Christensen 2004; Hedden 2013; Pettigrew 2019) stress, Dutch Book arguments standardly make use of the following thesis about fair prices:

> **Ramsey's Thesis (RT)**
>
> "Suppose your credence in X is p. Consider a \$$S$ bet on X [that pays \$$S$ if X and \$0 otherwise]... You are rationally required to pay \$$x$ for this bet, if $x < pS$." (Pettigrew 2019: 4)

Thus, for example, a probabilistically incoherent agent who assigns X credence .6 and Y credence .6 (when Y is logically equivalent to $\neg X$) is required to pay \$.55 for a bet that pays \$1 if X and to pay \$.55 for a bet that pays \$1 if Y. Taken together, these bets logically guarantee a loss of \$.10.

	Y	X
Bet 1	−\$.55	+\$.45
Bet 2	+\$.45	−\$.55

Given RT, it is natural to think of credence as a special kind of conditional attitude: having a .6 credence in X means being (inter alia) *logically committed to regarding Bet 1 as acceptable if it is offered for \$.55*. Having a .6 credence in X *and* a .6 credence in Y means being logically committed to regarding *both* Bets 1 and 2 as acceptable if they are each offered for \$.55. And so having a .6 credence in X *and* a .6 credence in Y means being logically committed to regarding as preferable something that one is also logically committed to not regarding as preferable, namely, the conjunction of Bets 1 and 2.

[31] See (a.o.) Rothschild (2012); Yalcin (2011, 2012a); Moss (2015, 2018).

[32] For an Accuracy-Dominance argument for probabilism, see Joyce (1998). For an Accuracy-Dominance argument for conditionalization, see Briggs & Pettigrew (2020). For an overview of Dutch Book arguments for probabilism, see Hájek (2009). For a Dutch Book argument for conditionalization, see Lewis (1999).

The bounded theory of the previous section *does not* account for this type of logical commitment (since it is meant to attach or apply to incoherent or unreasonable states of decision).[33] But it is easy to modify the theory so that it does: simply drop the stipulation that a decision state must be reasonable and/or coherent (e.g., by allowing decision states to be constructed with credal measures Pr such that $\Pr(X) + \Pr(\neg X) \neq 1$, and, more generally, to be unreasonable in any of the ways implied above).

For an illustration of concept, consider again a Dutch Book. On the relaxed version of this theory (and in contrast with the theory of the prior section), it would be possible for (the credal component of) a decision state to satisfy both:

(22) a. There's a 60% chance that X.
b. There's a 60% chance that Y.

Reading RT as a thesis about the logical commitments of states of partial belief (and spotting myself some details), any decision state that satisfies both of these sentences will also satisfy:

(23) a. If I'm offered Bet 1 for \$.55, it's preferable to take it.
b. If I'm offered Bet 2 for \$.55, it's preferable to take it.

Thus, supposing you are offered both Bets 1 and 2, you are (on the assumption that you are logically committed to reasoning via modus ponens) logically committed to concluding:

(24) a. It is preferable to take Bet 1 for \$.55.
b. It is preferable to take Bet 2 for \$.55.

So you regard it as preferable to take both bets.[34] But of course this "contradicts" another preference of yours: since the payout of both bets is negative, you in fact prefer taking *neither* bet to taking both. And so, given this "proof-theoretic" understanding of logical commitment, any decision state that satisfies (22) is logically committed to a contradiction-in-preference—one we could gloss by saying that, relative to that state, it is and is not preferable to take both bets if you are offered both.

An alternative strategy for theorizing Dutch Books is to say they, just like our original case (1), involve a kind of intuitive failure of modus ponens, in the following sense. If the relation of logical commitment is closed under modus ponens, then whenever an agent is logically committed to (23) and is offered both bets, they are logically committed to a conclusion *they are committed to evaluating as false*, namely, that it is better to take both bets (and lose a dime) than to reject both bets (and keep the dime).[35] Imagine, to the extent that you are able, that you had evidence that suggested both that it was probable that X and that it was probable that Y. Intuitively speaking, although this evidence would be (by assumption) reason to accept (23), it *is not* reason to accept that the package consisting of both bets is superior to the package consisting of neither.[36]

[33] There is also, to be sure, a fairly clear sense in which this sort of agent shouldn't be regarded as logically committed to this conclusion. An agent in this decision situation *makes a logical error* if she is offered both bets and decides to accept both (given that both bets logically guarantee a loss of \$.10).

[34] To see this more clearly, consider a standard assumption about the logic of strict preference, known as conjunctive expansion, according to which p is preferred to q iff $(p \wedge \neg q)$ is preferred to $(\neg p \wedge q)$. Note first that the agent in our Dutch Book is logically committed to preferring taking Bet 1 over rejecting Bet 2. Given conjunctive expansion taking both Bets 1 and 2 is preferred to rejecting both Bet 1 and Bet 2 iff taking Bet 1 is preferred to rejecting Bet 2.

[35] The notion of being "logically committed" to a conclusion is perhaps something of a misnomer for the somewhat more permissive notion I mean to invoke here. On the rough notion I have in mind, being logically committed to ϕ amounts to *being in a position to conclude* ϕ (compare Pettigrew 2019). Thus, an agent in a Dutch Book is in a position to conclude something they are also in a position to reject, namely, that the package consisting of both bets is superior to the package consisting of neither. I take the incoherence of this sort of position as given.

[36] A similar phenomenon is identified in work on "wide" and "narrow scope" normative requirements (see e.g. Broome 1999; Kolodny 2005). There is a sense in which believing $\neg\neg p$ supports believing p: if you believe $\neg\neg p$, you ought to believe p. But suppose you believe $\neg\neg p$ for *no good reason*. Does it follow (as it seems it should, by modus ponens) that you ought to believe p? Intuitively it does not: indeed, you ought not believe $\neg\neg p$, and since not believing $\neg\neg p$ logically commits you to not believing p, you *ought not* believe p. My best sense is that there are simply two notions of logical commitment at issue in this debate (which parallel the two notions of logical commitment developed in this section and the last). One notion is silent about the logical commitments of unreasonable states of opinion; the other notion is not. On the latter notion, you are logically committed to believing p, supposing you irrationally believe $\neg\neg p$; on the former notion, you are not.

Nevertheless, if logical commitment is closed under modus ponens (and can "attach" to representations of unreasonable or incoherent states of mind), then any decision state that accepts (23) is a decision state that is logically committed to this absurd conclusion. On this understanding of logical commitment, the same would hold for cases (1) and (19): conditional on an unreasonable or incoherent way of representing, an otherwise logically consistent agent is, in a predictable range of contexts, logically committed to a conclusion they would not accept (e.g., an absurdity or contradiction).[37]

I suspect this is the very notion of logical commitment philosophers employ when they argue that agents whose credences violate the laws of probability (or which they fail to update by conditionalization) are logically committed to accepting conclusions that are unacceptable from their own point of view.[38] On this notion of logical commitment, we would expect "intuitive" failures of modus ponens of this sort to run rampant: when the premises of a modus ponens argument correspond (in context) to an unreasonable way of representing, taking logical commitment to be closed under modus ponens will, in a predictable range of contexts, commit an (otherwise logically consistent) agent to reasoning from premises they would accept to premises they would reject.

4.4 Plural Commitment

I have no reason to reject such a notion, which seems to me necessary for running Dutch Book arguments as standardly understood (and exceedingly useful for running Dutch Book-type arguments to show that unreasonable or incoherent states of decision are mired in logical commitments that are unacceptable from their own point of view). I do, however, want to end with a brief for the type of "bounded" theory developed in the prior section.

A logic of decision, as I use the phrase, can be understood to provide two different types of standards for the normative assessment of deductive reasoning in a context. One kind of standard is "narrowly" logical (we might also say "syntactic" or "computational"): concluding ϕ from Γ is permitted in c just when for some decision state Π made salient in c, $\Pi \vDash \Gamma$, and $\Gamma \vdash \phi$. Another kind of standard is of a more "substantive" character: concluding ϕ from Γ is permitted in a context c just when for some decision state Π made salient in c, Π is reasonable in c, $\Pi \vDash \Gamma$, and $\Gamma \vdash \phi$. Since the latter kind of standard invokes a normative notion of reasonability, we would expect the intuitive applicability of the second kind of standard to vary with one's background conception of a reasonable state of decision.

This expectation is not disappointed, as I'll illustrate with the Newcomb Problem. There are two boxes before you, A and B. You keep what cash is inside any box you open. At the time of your decision, an exceptionally reliable predictor has already acted as follows:

- It has put $1,000 in A.
- It has predicted whether you will open A in addition to B.
- If it has predicted you will take one box (that you will not open A), it has put $1,000,000 in B.
- If it has predicted you will take two boxes (that you will open A in addition to B), it has put nothing in B.

The decision situation can be represented with the following table:

	Predicted one-boxing	Predicted two-boxing
One-box	$1,000,000	$0
Two-box	$1,001,000	$1,000

There is a notorious disagreement between Causal and Evidential Decision Theorists about the permissibility of dominance reasoning in the Newcomb Problem: EDT rejects it (since EDT permits dominance

[37]Decision theorists have puzzled over how to motivate the restriction to well-formed decision problems that characterizes a decision theory like Savage (1972), sometimes deriding them as ad hoc or external (see, e.g., Joyce 1999: Ch. 4). I have here provided what amounts to a Dutch Book argument for this restriction.

[38]Dutch Book arguments are often (mis-)read as supplying a *prudential* reason to conform one's credences to the laws of probability: roughly, if you don't so-conform, you're liable to become a money pump. This reading is impluasible, as Christensen (2004) emphasizes. It is true that Dutch Book arguments do not provide an *epistemic* (i.e., truth-related) reason to conform one's credences to the laws of probability, but they do nevertheless establish that agents whose credences violate the laws of probabilities are committed to certain forms of logical inconsistency or incoherence.

reasoning only in decision problems where the actions do not provide any evidence about which state is actual, which does not hold in Newcomb), CDT embraces it (since CDT permits dominance reasoning in decision problems where the actions are causally independent of the relevant states, as they are in the Newcomb Problem).

There should, however, be no disagreement about the acceptability of the following, *relative to* or *against* this decision situation. (Here read 'it is preferable to take two' simply as 'I get more money by taking two'.)

(25) a. If the predictor predicted I'd take one box, it is preferable to take two.
b. If the predictor predicted I'd I take two boxes, it is preferable to take two.
c. If the predictor predicted I'd take one box or it predicted I'd take two boxes, it is preferable to take two.
d. ?So, it is preferable to take two boxes.

CDT-ers conclude from this set of premises that it is (unconditionally) preferable to take both boxes, while EDT-ers reject the inference altogether.[39] The central difficulty of EDT's position on the Newcomb Problem, of course, is that the CDT-er's dominance argument is *facially* or *intuitively sound* (as assessed from the natural representation of the decision situation represented in the above decision table).[40]

Here I will sketch a strategy that allows the EDT-er to resist this assessment (by invoking the bounded notion of logical commitment described above).[41] By EDT's lights, this natural representation of the decision situation violates act-state (evidential) independence. For now-familiar reasons, I have said that decision states witnessing violations of act-state (evidential) independence deductively commit an agent in such a state to conclusions that are unacceptable (by EDT's lights), to wit, that it is preferable to take two boxes (and almost certainly walk away with $1,000) than to take one box (and almost certainly walk away with $1,000,000). The proponent of EDT should concede that (25c) is satisfied *when evaluated against such a decision state* (and that such a decision state is "narrowly" logically committed, by modus ponens, to a preference for taking two boxes). A way for the EDT-er to resist this argument's claim to *soundness* is to say that:

- The representation of the decision situation that is made salient in the Newcomb Problem is unreasonable (since the states evidentially depend on the actions).[42]
- There is a notion of satisfaction according to which it is not the case that (25c) satisfied in the Newcomb Problem, even though the relevant decision state meets the condition on decision states associated with (25c).

EDT can explain the pull of CDT's dominance argument in the Newcomb Problem as follows: although (25c) is strictly speaking unacceptable, it is smoothly evaluated as acceptable relative to a decision state made salient in the Newcomb Problem (that is to say, it is satisfied relative to this salient decision state).

This is *not*, as I'll explain, to say that (25c) is judged *false* in the context of use associated with the Newcomb Problem. The context of use makes salient a decision state that is (narrowly) logically committed to (25c), and so there is a sense in which (25c), like (1c), is simply evaluated as true in its context of use. The proponent of EDT can agree with *all of this* and still consistently maintain that they are not logically committed to a preference for two-boxing, provided they invoke a "substantive" notion of logical commitment, on which logical commitment does not attach to unreasonable states of decision.

[39] Indeed, EDT-ers commonly offer a Dutch Book-inspired argument against the pattern of reasoning recommended by CDT: roughly, if you're so smart why ain't you rich? The thrust of the EDT-er's argument is that a logic of decision should favor a course of action that will, with a high degree of probability, make you a millionaire over an action that will, with a high degree of probability, net you only $1,000.

[40] The intuitive soundness of this sort of dominance argument is surely a large part of the reason that authors like Lewis (1981); Egan (2007) (a.o.) are inclined to treat the Newcomb Problem as a *counterexample* to EDT.

[41] Jeffrey (1983)'s formulation of EDT is celebrated for abandoning the requirement of Savage (1972) that a decision state must be well-formed. This discussion can be read as suggesting that the EDT-er has reasons (akin to the reason provided by the Dutch Book argument for Probabilism) for invoking a notion of well-formed-ness in their own theory of deductive reasoning.

[42] Worth noting: from the point of view of EDT, such a state is committed to an inconsistency (on the assumption that the state also endorses EDT's preference for taking one box over two). For the proponent of EDT, it is a kind of *absurd state*. In my judgment there is nothing theoretically remarkable about a theory of satisfaction on which only non-absurd states stand in the satisfaction relation to a sentence.

4.5 *Judgments of Acceptability and Soundness*

I will end with a stab at a theory of *minimal* and *substantive* support/acceptance in context for preference statements of the form $\star\phi$ (which I will use to state a theory of soundness in context for deductive arguments). Let c be a context of use that makes salient a decision state (or, possibly, a plurality of decision states). Let $\vDash$ be the unbounded (total) satisfaction relation, and let $\vDash_T$ be a satisfaction relation that is defined only for decision states witnessing act-state independence by the lights of decision theory T. Then:

Definition 11. *$\star\phi$ is minimally accepted in c iff for any c-relevant Π, $\Pi \vDash \star\phi$.*[43]

Definition 12. *$\star\phi$ is T-accepted in c iff for any c-relevant Π, $\Pi \vDash_T \star\phi$.*

It is obvious that minimal acceptability is *not* closed under modus ponens. A simple counterexample: the decision state Π made salient in (1) is such that $\Pi \vDash$ (1c). But, by assumption, $\Pi \vDash \neg$(1d). On the assumption that (narrow) logical commitment is closed under modus ponens, Π is (narrowly) committed to (1d). In sum, Π is (narrowly) committed to ϕ and $\Pi \vDash \neg\phi$. Thus, Π is incoherent (on the assumption that narrow logical commitment is closed under modus ponens). (Note that it is essential to this model of Π's incoherence that $\vDash$ *not be* closed under modus ponens.)

While judgments of minimal acceptability are constrained solely by facts about which decision state(s) is relevant or salient in a context of utterance, verdicts of T-acceptability are additionally constrained by the notion of reasonability logically appropriate (given an unrestricted deductive "permission slip" to reason by modus ponens) to T. (25c) is minimally acceptable in the context of the Newcomb Problem. It is also acceptable "for" CDT. But it is not acceptable (nor is its negation acceptable) in context "for" EDT. That is not because (25c) has one satisfaction-condition in the mouth of the Evidential Decision Theorist and another in the mouth of the Causal Decision Theorist—on the present theory, it doesn't. It is instead because EDT does not regard the decision state made salient in the Newcomb Problem as a decision state from which an agent can safely reason deductively, i.e., by dominance!

If the acceptability of an argument from Γ to ϕ in c were understood as the conjunction of the minimal acceptability of Γ in c with the fact that $\Gamma \vDash \phi$, then (1), as well as (25), would be acceptable (i.e., it would be adjudged as *sound* relative to its context of use). Since (1) is unsound, acceptability is not to be understood in this fashion. And, since the soundness of (25) appears to depend on whether deductive dominance reasoning is permitted in the Newcomb Problem, there is some reason to think that intuitive judgments of soundness are to be regarded as, in this sense, normatively laden (and to construct a logic of decision that represents how such judgments are liable to shift with a background, possibly implicit or tacitly accepted, decision theory).

Like (25c), (1c) is minimally accepted in its context of use. Unlike (25c), (1c) is unratifiable in its context of use, given the following notion of ratifiability.

Definition 13. *Given a set of decision rules $\mathcal{T}$, $\star\phi$ is ratifiable in c iff, for some $T \in \mathcal{T}$, $\star\phi$ is T-true in c.*

Although (1c) is evaluated as acceptable relative to the salient decision state, *no* decision theory can permit dominance reasoning in such a decision state. So (1c), unlike (25c), is *unratifiable* in its context of use. This is meant to dispel the feeling of paradox around (1): the arguments I provided for the truth of (1c) were arguments for its minimal truth in this context. But although minimal truth is evidently enough to support an intuition of acceptability in context, it does not show that (1c) is to be regarded as acceptable—*as a premise from which one is permitted to reason deductively*—from the vantage point of *any* plausible computational rule for rational decision making. The argument from (1c) to (1d) should be judged unsound in its context, whatever one's decision rule.

4.6 *Conclusion*

While this paper considered several other strategies for dealing with the puzzle posed by (1), all appeared, in one way or another, to under-generate good deductive inferences. In lieu of retreading these points

[43] The quantifier 'for any c-relevant Π' should be read as requiring a non-empty domain: if no Π is salient or relevant in c, then it is not the case that $\star\phi$ is minimally accepted in c.

here, I'll end with some general comments on the understanding of *good deductive inference* appropriate to the strategy I've proposed.

The basic puzzle with which this paper grappled was this: the context I described for (1) appeared to make available the information in (1c), without seeming to make available the information in (1d). This suggests that a theory of good deductive inference in the following vein is false:

- A deductive inference from Γ to ϕ is good in c iff c makes available the information in Γ and $\Gamma \vdash \phi$.

Good deductive inference looks instead more like this:

- Given a normative theory T, a deductive inference from Γ to ϕ is good in c iff c makes available a T-reasonable way of representing the information in Γ and $\Gamma \vdash \phi$.

On this understanding, good deductive inferences preserve a property we could call *reasonably accepting as a premise for the sake of deductive reasoning*. On this way of thinking, the notion of good deductive inference is fundamentally epistemic, rather than doxastic or informational, in character.

References

Alonso-Ovalle, Luis. 2006. Disjunction in alternative semantics. http://semanticsarchive.net/Archive/TVkY2ZlM/alonso-ovalle2006.pdf. Ph.D. Dissertation, University of Massachusetts Amherst.

Ayer, A. J. 1964. Fatalism. In *The Concept of a Person and Other Essays*, 235–268. Macmillan & Co.

Biezma, María & Kyle Rawlins. 2012. Responding to alternative and polar questions. *Linguistics and Philosophy* 35: 361–406. doi:10.1007/s10988-012-9123-z.

Bledin, Justin. 2014. Logic informed. *Mind* 123: 277–316. doi:10.1093/mind/fzu073.

Bledin, Justin. 2020. Fatalism and the logic of unconditionals. *Noûs* 54: 126–161. doi:10.1111/nous.12257.

Briggs, R. A. & Richard Pettigrew. 2020. An Accuracy-Dominance argument for conditionalization. *Noûs* 54(1): 162–181. doi:10.1111/nous.12258.

Broome, John. 1999. Normative requirements. *Ratio* 12: 398–419. doi:10.1111/1467-9329.00101.

Buller, David J. 1995. On the 'standard' argument for fatalism. *Philosophical Papers* 24: 111–125. doi:10.1080/05568649509506524.

Cantwell, John. 2006. The logic of dominance reasoning. *Journal of Philosophical Logic* 35: 41–63. http://www.jstor.org/stable/30226858.

Cariani, Fabrizio, Magdalena Kaufmann & Stefan Kaufmann. 2013. Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy* 36: 225–259. doi:10.1007/s10988-013-9134-4.

Charlow, Nate. 2013b. What we know and what to do. *Synthese* 190: 2291–2323. doi:10.1007/s11229-011-9974-9.

Charlow, Nate. 2013c. Conditional preferences and practical conditionals. *Linguistics and Philosophy* 36: 463–511. doi:10.1007/s10988-013-9143-3.

Charlow, Nate. 2015. Prospects for an expressivist theory of meaning. *Philosophers' Imprint* 15: 1–43.

Charlow, Nate. 2016. Decision theory: Yes! Truth conditions: No! In N. Charlow & M. Chrisman (eds.) *Deontic Modality*, 47–81. Oxford University Press.

Charlow, Nate. 2018. Decision-theoretic relativity in deontic modality. *Linguistics and Philosophy* 41: 251–287. doi:10.1007/s10988-017-9211-1.

Christensen, David. 2004. *Putting Logic in its Place: Formal Constraints on Rational Belief.* Oxford: Oxford University Press.

Ciardelli, Ivano. 2018. Questions as information types. *Synthese* 195: 321–365. doi:10.1007/s11229-016-1221-y.

Ciardelli, Ivano & Floris Roelofsen. 2011. Inquisitive logic. *Journal of Philosophical Logic* 40: 55–94. doi:10.1007/s10992-010-9142-6.

Dreier, James. 2009. Practical conditionals. In David Sobel & Steven Wall (eds.) *Reasons for Action*, 116–133. Cambridge: Cambridge University Press.

Dummett, Michael. 1964. Bringing about the past. *Philosophical Review* 73: 338–359. doi:10.2307/2183661.

Egan, Andy. 2007. Some counterexamples to causal decision theory. *The Philosophical review* 116: 93–114. http://www.jstor.org/stable/20446939.

von Fintel, Kai. 1999. NPI licensing, strawson entailment, and context dependency. *Journal of Semantics* 16: 97–148. doi:10.1093/jos/16.2.97.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press.

Gibbard, Allan & William L. Harper. 1981. Counterfactuals and two kinds of expected utility. In W. Harper, R. Stalnaker & G. Pearce (eds.) *Conditionals, Belief, Decision, Chance and Time*, 153–190. Springer Netherlands.

Gillies, Anthony S. 2009. On truth-conditions for *If* (but not quite only *If*). *The Philosophical Review* 118: 325–349. doi:10.1215/00318108-2009-002.

Gillies, Anthony S. 2010. Iffiness. *Semantics and Pragmatics* 3: 1–42. doi:10.3765/sp.3.4.

Groenendijk, Jeroen & Floris Roelofsen. 2009. Inquisitive semantics and pragmatics. http://www.illc.uva.nl/inquisitive-semantics. Ms., ILLC.

Groenendijk, Jeroen & Martin Stokhof. 1984. Studies on the semantics of questions and the pragmatics of answers. http://dare.uva.nl/record/123669. Ph.D. Dissertation, University of Amsterdam.

Hájek, Alan. 2009. Dutch book arguments. In P. Anand, P. Pattanaik & C. Puppe (eds.) *Oxford Handbook of Rational and Social Choice*, 173–195. Oxford: Oxford University Press.

Hamblin, Charles. 1958. Questions. *Australasian Journal of Philosophy* 36: 159–168. doi:10.1080/00048405885200211.

Hedden, Brian. 2013. Incoherence without exploitability. *Noûs* 47: 482–495.

Hospers, John. 1967. *An Introduction to Philosophical Analysis*. London: Routledge.

Jeffrey, Richard. 1983. *The Logic of Decision*. Chicago: University of Chicago Press, 2nd edn.

Joyce, James. 1998. A nonpragmatic vindication of probabilism. *Philosophy of Science* 65: 575–603. doi:10.1086/392661.

Joyce, James. 1999. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Kolodny, Niko. 2005. Why be rational. *Mind* 114: 509–563. doi:10.1093/mind/fzi509.

Kolodny, Niko & John MacFarlane. 2010. Ifs and oughts. *Journal of Philosophy* 107: 115–143. doi:10.5840/jphil2010107310.

Kratzer, Angelika. 1981. The notional category of modality. In H. Eikmeyer & H. Rieser (eds.) *Words, Worlds, and Contexts*, 38–74. Berlin: De Gruyter.

Kratzer, Angelika. 1991. Conditionals. In A. von Stechow & D. Wunderlich (eds.) *Semantics: An International Handbook of Contemporary Research*, 651–656. Berlin: De Gruyter.

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

Lewis, David. 1981. Causal decision theory. *Australasian journal of philosophy* 59: 5–30. doi:10.1080/00048408112340011.

Lewis, David. 1999. Why conditionalize? In *Papers in Metaphysics and Epistemology*, 403–407. Cambridge University Press. doi:10.1017/CBO9780511625343.024.

Moss, Sarah. 2015. On the semantics and pragmatics of epistemic modals. *Semantics and Pragmatics* 8: 1–81.

Moss, Sarah. 2018. *Probabilistic Knowledge*. Oxford: Oxford University Press.

Pettigrew, Richard. 2019. On the expected utility objection to the dutch book argument for probabilism. *Noûs* doi:10.1111/nous.12286.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In *OSU Working Papers in Linguistics, Vol 49: Papers in Semantics*. Ohio State University. http://semanticsarchive.net/Archive/WYzOTRkO/.

Rothschild, Daniel. 2012. Expressing credences. *Proceedings of the Aristotelian Society* CXII, Part 1: 99–114. doi:10.1111/j.1467-9264.2012.00327.x.

Savage, Leonard J. 1972. *The Foundations of Statistics*. Dover.

Schulz, Moritz. 2018. Modus ponens under the restrictor view. *Journal of Philosophical Logic* 47: 1001–1028. doi:10.1007/s10992-018-9459-0.

Stalnaker, Robert. 1975. Indicative conditionals. *Philosophia* 5: 269–286. doi:10.1007/BF02379021.

Stojnić, Una. 2017. One's modus ponens: Modality, coherence and logic. *Philosophy and Phenomenological Research* 95: 167–214. doi:10.1111/phpr.12307.

Veltman, Frank. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25: 221–261. doi:10.1007/BF00248150.

Willer, Malte. 2012. A remark on iffy oughts. *Journal of Philosophy* 109: 449–461. doi:10.5840/jphil2012109719.

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

Yalcin, Seth. 2012a. Bayesian expressivism. *Proceedings of the Aristotelian Society* CXII, Part 2: 123–160. doi:10.1111/j.1467-9264.2012.00329.x.

Yalcin, Seth. 2012b. A counterexample to modus tollens. *Journal of Philosophical Logic* 41: 1001–1024. doi:10.1007/s10992-012-9228-4.