2

Decision Theory: Yes! Truth Conditions: No!

Nate Charlow

1 Introduction

This essay attempts to leverage a special theoretical pressure involved in accounting for the meaning and function of language with loosely-speaking *normative* or *action-guiding* meaning. A core case—and my focus here—is that of deontic modals such as *should* and *ought*.[1] Many authors claim that a "standard" semantics for deontic modals—on which they express varieties of quantification over contextually determined domains of deontically admissible possibilities—is unviable. In particular, the standard semantics is said to ignore—in virtue of lacking the structure to represent—factors relevant to determining the truth of a deontic modal in a given situation. To yield plausible truth-verdicts for claims expressing what should or may be done in various kinds of situation, a semantics for such claims should enrich itself with the fruits of the normative literature on how to determine and reason about what should or may be done in such situations. By and large, existing accounts suggest the way to fix this is by *adding structure*—much implicitly borrowed from normative theory—to the standard semantics.

The literature has grown thick with proposals in this vein (and with theoretical and empirical justifications for complications that would otherwise appear unparsimonious). But it is comparatively thin on methodological and philosophical reflection,

[1] Linguists used to standard ways of carving up modals in natural language will think "deontic modal" a poor name for the focus of this chapter. Better would be "prioritizing modal," a label used to cover modals with a broadly action-guiding or advisory function (namely: proper deontic, bouletic, and teleological) (Portner 2007). Philosophers tend to use "deontic modal" to cover roughly the same ground as the linguist's "prioritizing modal." Although I tend to prefer "practical modal" here (see Charlow 2013a), I follow the philosophers here. (Most of the chapters of this volume do so too.) Allow me to also note that my primary focus in this chapter is on so-called *weak* deontic necessity modals (such as *should* and *ought*) whose content is directly paraphrasable in terms of comparative betterness (on which more below; cf. especially Kratzer (1981) on "weak necessity"). *Strong* deontic necessity modals (for example, *must*) are a more difficult case. (Like most who have written on the distinction, I am inclined to treat strong deontic necessity as a special case of weak deontic necessity.)

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 48 — #2

whether about the endeavor itself, where the significant theoretical divisions and methodological defaults lie,[2] or the principles—besides the constraint that requires a correct semantics to match considered judgments of correctness and consequence—that might govern this sort of theorizing. *Should* we augment our theorizing about the semantics of deontic modals with the fruits of normative theory? If so, *how* should we do it? Have past efforts gone wrong in important ways?

This chapter tries to address such questions. It first attempts to build the strongest possible case for packing the fruits of the normative literature into the semantics of deontic modals. Unfortunately, the case crumbles on examination. There are metasemantic reasons to demand that semantics *remain neutral* on certain normative matters—namely, all of them. We get closer to the truth about the semantics of deontic modals by simply *excising* normative assumptions from our semantic theory. This, however, seems to mean giving up on the project of matching the verdict of truth delivered by a semantics to the one delivered by intuition (or so I will argue).

This is a pessimistic conclusion, but this chapter has an optimistic response in mind: a view that might be clumsily called *Nonpropositionalist, Minimalist Expressivism*. I will suggest that deontic modals support a reorientation of semantic theorizing, away from the aim of delivering an account of a sentence's truth-condition, toward the aim of modeling the state of competently accepting a sentence ("*Nonpropositionalism*"). More precisely: the correct semantics for deontic modals is one that lays bare the manifold ways in which a normative view can be combined with linguistic competence with respect to deontic modals to generate a concrete linguistic judgment of the form *A should do X*. The core compositional semantics for deontic modals is exhausted by the idea that deontic modals quantify over a domain determined by a sequence of parameters (of basically any length) whose values are fixed by the context of use (namely, by the specific normative view of a subject) ("*Minimalism*"). The project of attempting to understand the "content" or "truth-condition" of a deontic modal on an occasion of use by investigating its compositional semantics should, in large part, be abandoned. (This, I argue, appears less of a cost when married to the right pragmatics—an "*Expressivist*" pragmatics.)

What is the upshot, if I am right? Corresponding issues clearly arise for language with broadly normative or action-guiding or advisory meaning—moral language, imperatives, and so forth. If I am right, much of the work in these areas needs to be reconsidered, on the grounds that its authors are engaged in normative theorizing, rather than semantics. How far this kind of critique can be extended to recent, "cognitively laden" work on, e.g., modals, probability operators, and conditionals is a question I will not address here, but which will merit future attention.

[2] A vignette: the type of amendment to Kratzer's view suggested in, for example, Charlow (2013b); Cariani (2013) is variously portrayed as a "challenge" and a "conservative amendment" to the "classical semantics" (see respectively von Fintel 2012; Carr 2012a). This is a theoretically confounding state of affairs (about which I will have more to say below). (Let me note, however, that Carr's work is an increasingly influential exception to the trend I am referencing here.)

2 The Standard Semantics

What does a sentence like (1), which I will refer to as a *deontic necessity modal* (DNM), say?

(1) You should wash your hands.

In one sense, we know pretty well: (1) says roughly that *washing your hands is better (given the relevant information and body of standards) than not* (see especially Kratzer 1981, 1991). To see why these claims are approximately equivalent, here is an informal proof. First, suppose you should wash your hands, and suppose for reductio it is *not* the case that washing your hands is better than not. Then not washing your hands is at least as good as washing your hands (or they are incomparable). In either case, it hardly seems in any sense obligatory that you wash your hands; so, contrary to supposition, it is not the case that you should wash them. In the other direction, suppose washing your hands is better than not and suppose for reductio that it is not the case that you should wash your hands. Then it would seem permissible for you not to wash your hands. This cannot be the case if washing your hands is better than not: something is optional only if not outranked by the alternatives.

While this apparently innocent gloss on the meaning of DNMs hardly seems likely to tell us anything of interest about their semantics, it in fact leads more or less directly to a substantive semantics for them. It leads, namely, to the familiar story on which DNMs express universal quantification (or something close enough) over a domain of best possibilities. I will present the case for this formally, but, for those inclined to skip ahead, the basic idea is this: given some relatively innocent assumptions about the relevant *better than* relation, p is better than $\neg p$ iff seeing to it that $\neg p$ yields a sub-optimal outcome, iff all of the best (i.e. good-enough) possibilities are possibilities in which p. The remainder of this section firms up this gloss while introducing a bit of apparatus that will be used at various points in the chapter; it can be skipped without much harm.

To begin, notice that our rough gloss seems to involve just two core notions. First, *better than*, indicating ranking according to some kind of preference or standard. Second, *than not*, indicating that the ranking applies to some range of relevant options or alternatives. The oldest (and still standard) way of modeling these notions in linguistic semantics understands them *modally* (Kratzer 1977, 1981, 1991; Lewis 1974, 1979b). Begin with a domain of relevant alternative possibilities R (*than not*),[3] ordinarily typed as a set of possible worlds,[4] commonly called a Modal Base.

[3] I will remain silent on what it takes for a possibility to be a relevant alternative when interpreting a DNM.

[4] We will understand possible worlds as functions from atomic sentences to truth-values—i.e., as something that settles whether grass is green, sheep scream, etc. It is open to us to understand it as a set of more complicated entities—worlds centered on a time coordinate (or agent or whatever), relations on worlds (i.e., actions), and so on. For a sense of the different options here, see Cariani (2013); Cariani et al. (2013); Charlow (2009, §4).

Definition 1. A **modal base** R is a set of relevant possibilities.

Next, we must rank the stuff in R according to some criteria (*better*). Usually possibilities are ranked relative to a weak preorder $\preceq$ on R, where '$w \preceq v$' means approximately that w is as least as good as v. The weak preorder is usually assumed to obey certain axioms (analogous to the axioms of order governing preference orders in decision theory): at minimum, it is Reflexive ($w \preceq w$) and Transitive ($w \preceq v$ and $v \preceq u$ imply $w \preceq u$).[5] Such a ranking is commonly called an Ordering.

Definition 2. An **ordering** $\preceq$ on R is a weak preorder on R that is Reflexive and Transitive.

From this weak preorder, a strict order $\prec$ may be defined:

Definition 3. The **strict ordering** $\prec$ is the relation w bears to v iff $w \preceq v$ and $v \npreceq w$.

Note: since $\prec$ is by definition Asymmetric ($w \prec v$ implies $v \nprec w$) and Transitive, it is Acyclic: for no $w_1, \ldots, w_n \in R$ such that $w_1 \prec w_2, \ldots, w_{n-1} \prec w_n$ is it the case that $w_n \prec w_1$. It is easily lifted to propositions (sets of possibilities) as follows (see especially Kratzer 1981):

Definition 4. The **lifted ordering** $\preceq$ is the relation p ($p \subseteq W$) bears to q ($q \subseteq W$) iff, for all $v \in q$ there is some $w \in p$ such that $w \preceq v$.

Definition 5. The **lifted strict ordering** $\prec$ is the relation p bears to q iff $p \preceq q$ and $q \npreceq p$.

The definition most important for our purposes here is **Definition 5**. According to our innocent gloss, DNMs are interpreted relative to a Modal Base and Ordering. What conditions does a DNM express on a Modal Base and Ordering? In other words, under what conditions is a DNM $\Box p$ true at $\langle R, \preceq \rangle$? According to our gloss, $\Box p$ says that p is better than $\neg p$, relative to $\langle R, \preceq \rangle$. What does this mean? It would seem to mean that there's no $\neg p$ possibility ($\in R$) that cannot be improved on by some p-possibility (also $\in R$); p is better than $\neg p$ relative to $\langle R, \preceq \rangle$ just if, if $\neg p$, it is always possible to better (i.e., if p). Formally (letting $[\![\phi]\!]$ give the set of possibilities ϕ maps to true):

(Q1) $$\forall x \in [\![\neg p]\!] \cap R : \exists y \in [\![p]\!] \cap R : y \prec x$$

Informally: seeing to it that $\neg p$ will always leave things sub-optimal. It is easy to see p is better than $\neg p$, by the lights of Q1, iff $p \prec \neg p$, by the lights of **Definition 5**. Recall that $p \prec \neg p$ iff $p \preceq \neg p$ and $\neg p \npreceq p$. Hence, iff each $\neg p$ world is equaled or bettered by some p world, but not vice versa (so that at least one p-world betters any $\neg p$-world). In other words, it seems that a condition of the form $X \prec Y$ is a reasonable way of encoding the content of a claim of the form *X is better than Y*.

[5] Possibly, according to Kratzer, it is partial: for some w and v: $w \npreceq v$ and $v \npreceq w$. Partiality introduces complications that I do not wish to discuss in the main text—and its original motivation is frankly quite weak. I will implicitly assume that the ordering is non-partial whenever I need it to be total (see, e.g., the sort of case mentioned in Footnote 7).

Now suppose we make the Limit Assumption (Lewis 1973), so that there is always some non-empty set of $\preceq$-*best* worlds in R:[6]

(Limit) $$\exists x \in R : \neg\exists y \in R : y \prec x$$

. . .and we let $min_{\preceq}(R)$ designate the set of $\preceq$-*best* worlds in R.

Definition 6. $w \in min_{\preceq}(R)$ iff $\neg\exists v \in R : v \prec w$

On this assumption, a condition of the form $p \prec \neg p$ is equivalent to the following truth-condition, which says that the domain of sufficiently good possibilities is p-entailing:[7]

(Q2) $$min_{\preceq}(R) \subseteq [\![p]\!]$$

Here, finally, is an informal proof of the main "result" of this discussion:

$p \prec_R \neg p \;\Leftrightarrow\; min_{\preceq}(R) \subseteq [\![p]\!]$

$\Rightarrow$ Suppose $p \prec_R \neg p$. Suppose for reductio that for some $v \in min_{\preceq}(R) : v \in \neg p$. Then there is no world in R that betters v. This contradicts the first supposition.

$\Leftarrow$ Suppose $min_{\preceq}(R) \subseteq p$. Suppose for reductio that $p \not\prec_R \neg p$. Then either $p \not\preceq_R \neg p$ or $\neg p \preceq_R p$. Suppose $p \not\preceq_R \neg p$. Then some $\neg p$-worlds are not equaled/bettered by any p-world. Contradiction. So suppose $\neg p \preceq_R p$. Then each p-world is equaled/bettered by some $\neg p$-world. Contradiction. □

Signpost. The point of that was this: the Standard Semantics—on which deontic necessity involves universal quantification (and deontic possibility existential quantification) over $min_{\preceq}(R)$—seems to follow from a *very minimal* understanding of

[6] The Limit Assumption is reasonable insofar as the domain of relevant alternative possibilities should be presumed to be finite (for its negation implies that the domain is infinite). I think the Limit Assumption is reasonable, at least for the deontic case, and I will make it here. Nothing of ultimate importance will turn on this, however.

[7] I do not mean to suggest that this should be regarded as a revelation. However, standard reasons for thinking so are arguably mistaken. von Fintel and Iatridou (2008), for example, write: "Kratzer (1991) distinguishes between necessity and weak necessity as well. Her informal characterization is similar to ours: p is a weak necessity iff p is a better possibility than *not* p. The technical implementation is different from ours and crucially involves not accepting that there is always a set of most favored worlds (what is known as the Limit Assumption in the trade). *It appears to us that if one makes the Limit Assumption, Kratzer's definitions collapse, leaving no distinction between simple necessity and weak necessity*" (p. 118 n. 8, emphasis added). Contra the view expressed in this passage, the Limit Assumption is not sufficient for the relevant definitions to collapse. Consider a finite model (a fortiori satisfying the Limit Assumption) with $R = \{x, a, b\}$ in which $p = \{x, a\}$ and in which a and b are equally good ($a \preceq b$ and $b \preceq a$), but x is incomparable to both a and b. Obviously, since $a \preceq b$, $p \preceq \neg p$. But since $b \not\preceq x$, $\neg p \not\preceq p$. Hence $p \prec \neg p$. Obviously, however, $min_{\preceq}(R) = R$—since no world in R can be strictly improved on—hence $min_{\preceq}(R) \not\subseteq p$. The easiest way around this is to assume away the troublesome forms of partiality in $\preceq$, and I have implicitly done so in the main text. Thanks here to Fabrizio Cariani, Alex Silk, and Eric Swanson.

what deontic modals are in the business of saying (together with some relatively innocent assumptions about how to structure the formal semantic apparatus). This fact could, moreover, very reasonably be taken as a kind of roundabout vindication for the semantic value that, e.g., Kratzer (1981, 1991) assigns to necessity modals generally. On Kratzer's classic treatment, necessity modals exhibit something like polysemy. The necessity modal is assigned a context-independent, doubly-parametrized semantic value, that, in context, maps values for these two parameters—the Modal Base (analogous to our relevant alternative space) and Ordering Source (analogous to our Ordering) into a universally quantificational truth-condition. DNMs are distinguished from E(pistemic)NMs, not by quantificational force, rather by being interpreted in contexts in which deontic-specific assignments of the relevant pair-values are appropriate. The fact that an independently motivated Standard Semantics for DNMs can be derived from the Kratzer semantics for NMs is no small point in favor of the latter.

3 Refinement

It would be natural to draw the following moral from the foregoing: accounting for facts for which the Standard Semantics fails to account will mean *refining the basic apparatus of the Standard Semantics*, often times in ways that derive insight and plausibility from phenomena falling within the domain of practical rationality. Let us consider some examples.

3.1 Strong and weak DNMs

Notice that, given a fixed R and $\preceq$, the Standard Semantics has difficulty distinguishing the truth-condition of a "strong" DNM (*you must wash your hands*) and a "weak" DNM (*you should wash your hands*). (The standard story is that these must be truth-conditionally distinct, since strong DNMs apparently entail corresponding weak DNMs but not vice versa.) Thus von Fintel and Iatridou (2008) suggest weak DNMs are sensitive to (at least) two kinds of Ordering—one of which (Primary) ranks alternatives according to their realization of a "designated goal", the other of which (Secondary) further refines the domain generated by the Primary Ordering by ranking possibilities in it according to how well they meet various subsidiary considerations (e.g., cost or ease). On this view, the Standard Story is right for what we might call the *basic deontic modality*—the one whose interpretation is as sketched above—but must be refined before it can be extended to what we might call *secondary* or *derived deontic modalities*, such as weak DNMs, which have roughly the following as their truth-condition (with $\preceq_n$ the n^{th}-ary Ordering):

(Q3) $$min_{\preceq_2}(min_{\preceq_1}(R)) \subseteq [\![p]\!]$$

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 53 — #7

This refinement seems to derive plausibility from reflection on the parallel character of practical rationality—more precisely, on an aspect of practical rationality that the unrefined semantics seems to ignore. Treating some goal as *designated* in practical reasoning, as evoked in the following sort of instruction, means treating it as something that (in Dworkinian, Rawlsian, and Nozickian idioms, respectively) trumps, is lexically prior to, or functions as a side-constraint on the pursuit of secondary considerations. To illustrate, consider:

(2) Suppose my goal is to get to Harlem from Brooklyn. What should I do?

It is beside the point to note that you will get some enormous prize if you stay in Brooklyn; the practical question *what should I do?* queries which ways *of getting to Harlem* are best, while taking no note of rankings amongst the various ways of not getting to Harlem. Similarly, notice that, no matter how highly an alternative $w \in R$ ranks according to $\preceq_2$, if w does comparatively poorly by the metric of $\preceq_1$, so that $w \notin min_{\preceq_1}(R)$, then necessarily $w \notin min_{\preceq_2}(min_{\preceq_1}(R))$.

3.2 Epistemic and deontic orderings

Here is another illustration.[8] Consider the following impeccable practical inference.

(3) If you can wash your hands or can wipe them on your pants, you should wash them. $(if\, X \vee Y)(\Box A)$
If you can wash your hands or can use hand sanitizer, you should wash them. $(if\, X \vee Z)(\Box A)$
✓So, if you can wash your hands or can wipe them on your pants or can use hand sanitizer, you should wash them. $(if\, X \vee (Y \vee Z))(\Box A)$

(3) is *apparently* underwritten by the following truism of practical rationality: if p is preferable to both q and r, given a choice between p and $(q \vee r)$, p rationally must remain preferable. Relatedly, and in fact more generally, if p is preferable given all the relevant alternatives individually, p rationally must remain preferable given all the relevant alternatives collectively, as illustrated by, e.g., a case in which the relevant alternatives are given by the circumstances of the decision problem:

(4) If your fellow prisoner defects, you should defect. $(if\, X)(\Box A)$
If your fellow prisoner cooperates, you should defect. $(if\, Y)(\Box A)$
✓So, whatever your fellow prisoner does, you should defect. $(if\, X \vee Y)(\Box A)$

[8] Inspired by discussion in Lassiter (2011) of Kratzer's treatment of the *better possibility than* relation. A difference worth noting is that Lassiter is not worried about inferences involving modals and conditionals in the object language. He is rather worried about the fact that, regardless of whether *better* has a preferential or probabilistic meaning, Kratzer's treatment of the metalinguistic *better possibility than* relation validates the inference from (i) *A* is a better possibility than *B*, (ii) *A* is a better possibility than *C*, to (iii) *A* is a better possibility than *B* or *C*.

The reasoning of which this example is apparently an instance is apparently related to what is generally known as the *Statewise Dominance* constraint on choice functions in the literature on rational choice (which is related to the *Sure Thing Principle* of Savage 1972).[9]

Now consider a minimal variation of (3), which is rather less than impeccable. Note: *should* in (5) is meant to express a form of epistemic necessity concerning one's expectations about Al.

(5) If Al can wash his hands or can wipe them on his pants, he should wash them.
$(if\ X \vee Y)(\Box_{ep} A)$
If Al can wash his hands or can use hand sanitizer, he should wash them.
$(if\ X \vee Z)(\Box_{ep} A)$
#So, if Al can wash his hands or can wipe them on his pants or can use hand sanitizer, he should wash them.
$(if\ X \vee (Y \vee Z))(\Box_{ep} A)$

In (5), a formally similar inference fails, owing to something like a truism of epistemic rationality: even if p is considerably likelier than both q and r—sufficiently so to license the claim that if either p or q may be true, it should be p, and the claim that if either p or r may be true, it should be p—it doesn't follow that p is considerably likelier (to the same degree) than $(q \vee r)$.

Suppose with Kratzer (1981, 1991) (and most proponents of the Standard Semantics) that epistemic and deontic *should* express contents that are resolutions-in-context of a single parametrized semantic value—i.e., resolutions obtained by fixing, relative to a context of interpretation, values for parameters in the semantic value for *should*. This pair would seem strongly to suggest that the status of such inferences should follow not merely from facts about the parametrized semantic value of *should*. Instead, they should follow from these facts together with, on the one hand, stipulations about the character of the domain of quantification for modals expressing deontic necessity, as well as, on the other hand, stipulations about the character of the domain of quantification for modals expressing epistemic necessity.[10] Such stipulations, whatever shape they take, will apparently be underwritten by the above truisms of practical and epistemic rationality.

Signpost. The natural spin on the cases we considered in this section is that they motivate certain emendations to the Standard Semantics. More specifically, they seem

[9] The relationship between Statewise Dominance and the Sure Thing Principle is in fact complicated (as is the matter of their relationships to the examples mentioned here). Since none of the points I make will turn on the fact that it is the Sure Thing Principle, as opposed to Statewise Dominance (or indeed another dominance principle entirely), that is encoded within a semantics for DNMs, I will tend to speak loosely about these relationships in the main text.

[10] Lassiter (2011) draws the lesson that epistemic and deontic *should* probably do not share a common semantic core (and that these modals do not ultimately have a quantificational semantics). While I am sympathetic to the second part of his conclusion, it is several steps removed from the bare data we are considering here.

to highlight that the Standard Semantics is too coarse-grained to represent phenomena of normative significance which seem to bear clearly on the truth-conditions of DNMs. The way out is to refine the Standard Semantics to allow it to represent these sorts of normative phenomena, preliminarily along the lines suggested in this section. A large literature is devoted to pursuing a research program broadly in this vein.

4 Uncertainty

A key opportunity for refinement explored in the literature on DNMs concerns the semantic role of uncertainty. Specifically, the Standard Semantics has a hard time accommodating the manifest importance of uncertainty in determining what we should do. While some—Dowell (forthcoming); von Fintel (2012), not to mention Kratzer (2012)—claim the Standard Semantics can accommodate uncertainty as-is, it is commonly accepted that some sort of refinement is needed (as I'll explain). A number of people have attempted to provide one (see Carr 2012a; Cariani et al. 2013; Charlow 2013b; Silk 2014). Here I will review the major issues and build a case for, in Kratzer's phrase, "packing in" some of the fruits of decision theory into the semantics of DNMs.

4.1 MaxiMax

Here's a simple schematic case. Possibilities where g_3 is realized are preferred to possibilities in which g_2 is realized, which are in turn preferred to possibilities in which g_1 is realized:

$$g_3 \prec g_2 \prec g_1$$

Your available actions: *A* and *B*. The relevant contingency—the contingency that determines the respective outcomes of *A*-ing and *B*-ing—is whether *p*. *B*-ing brings a sure outcome in which g_2 is realized, while *A*-ing brings a good outcome (g_3) if *p*, but a relatively bad outcome (g_1) if $\neg p$.

	p	$\neg p$
A	g_3	g_1
B	g_2	g_2

In short, you prefer *A*-ing if *p* to *B*-ing, which you in turn prefer to *A*-ing if $\neg p$. Now: *what should you do* in this sort of decision problem? We cannot, it seems, answer this question without further information regarding, at least, the likelihood of *p* (as well as information about how preferable g_3 is to g_2 and how preferable g_2 is to g_1). Absent such information, what you should do is *indeterminate*.

The Standard Semantics—on a perhaps naïve reading (see §4.2)—seems, wrongly, to say that the answer is determinate. Possibilities in which g_3 is realized are strictly preferred to possibilities where it is not. The best possibilities (according to the relevant

ordering $\prec$) are thus all possibilities where g_3 is realized. But all the possibilities where g_3 is realized are possibilities where you A. And so all the best possibilities are possibilities where you A: $A \prec B$. As Carr puts it, the Standard Semantics seems to encode the decision rule MaxiMax. "This is a straightforward consequence of the more basic commitment...that we should always simply bring about the best possible outcome in the modal background" (Carr 2012a, p. 177). If we had to pick any decision rule for our semantics for DNMs to encode, MaxiMax would be *far* down the list.

4.2 Decision-theoretic quietism

I bring up this sort of case to introduce a possible response on behalf of the Standard Semantics before approaching a more well-known case. The response is familiar from the decision-theoretic dialectic surrounding, e.g., the Ellsberg Paradox: the decision table—more precisely, its preferential dimension—is said to be under-described. In the Ellsberg Paradox, the charge is that the payoffs explicitly represented in the decision problem ignore the role of *risk-aversion* in informing the preferences of apparently rational agents. Analogously, in this case, the charge might be that the relevant payoffs should be enriched. What is (dis-)valuable outstrips what can be represented with payoffs g_1–g_3. Also valuable is choosing an action that maximizes expected value (or a suitable stand-in for expected value). More generally, agents manifestly value choices that meet procedural conditions on how such choices are made (call such preferences *rational preferences*). Rational preferences manifestly bear on the desirability of actions that implement them (or fail to).[11]

Let's take this a bit further. In the schematic case in §4.1, it is natural to ascribe *first-order preferences*—preferences that supervene on your preferences among outcomes—of broad shape:

$$(A \wedge p) \prec_{FO} B \prec_{FO} (A \wedge \neg p)$$

Nevertheless, it is still clearly possible (for, e.g., an expected-utility maximizer) to have *considered preferences* among actions of shape:

$$B \prec_{ATC} A$$

Given **Definition 5** (and, indeed, any reasonable understanding of the properties of $\prec$) necessarily $\prec_{FO} \neq \prec_{ATC}$. So long, then, as we allow $\prec_{ATC}$, rather than $\prec_{FO}$, to determine the domain of quantification for the DNM, we can avoid the problem posed above.

This reply is decision-theoretically quietist, in the following sense. Contra what was suggested in §4.1, MaxiMax is in fact rejected as a "theorem" of the semantics.[12] However, *no replacement decision theory is encoded in its stead*. Indeed, it would seem

[11] Such an account is presented approvingly in an unpublished reply by Kratzer to Cariani et al. (2013), and echoed in von Fintel (2012). For criticisms similar to the one developed here, see Carr (2015); Silk (2014).

[12] Compare Cariani (2016): "A semantic theory might select the best worlds out of an ordering, but [this] does not mean that it implements MaxiMax."

that effectively any recipe for transforming $\prec_{FO}$ into $\prec_{ATC}$ must be *invisible to the compositional semantics of DNMs*.

Kratzer's unpublished argument for this sort of invisibility is a question whose answer is intended as obvious: "Why pack information about rational decision making into the meaning of modals?" A perhaps more perspicacious version of this question would be as follows: why pack any more information about rational decision making into the meaning of modals than we have to in order to make our semantics minimally predictive? After all, Kratzer's stipulations about the ordering $\preceq$ encode certain "axioms" of rational decision-making, including:

- Kratzer Dominance $(p \prec q \wedge p \prec r) \Rightarrow p \prec (q \vee r)$
- Acyclicity of $\prec$ $p_1 \prec \ldots \prec p_n \Rightarrow p_n \not\prec p_1$
- Independence of Irrelevant Alternatives (see Charlow 2013b):
  $(w \in min_{\preceq}(R) \wedge w \in R' \wedge R' \subseteq R) \Rightarrow w \in min_{\preceq}(R')$

Any and all object-language expressions of these features of the preference ordering are thus treated as *theorems* within Kratzer's semantics. These, however, are intuitively "good" things to treat as theorems within the relevant formal system (whereas it would be intuitively bad to render, e.g., a particular decision rule such as Expected Utility Maximization or MaxiMin a theorem).

Why pack any more into the compositional semantics of DNMs than we have to in order to generate the sorts of theorems we regard as desirable? I will refer to the proponents of the "obvious" answer to this rhetorical question as Decision-Theoretic Quietists. In the next section I will apply Quietism to a well-known puzzle case in the semantics of DNMs.

4.3 Applying decision-theoretic quietism: the miners

Kolodny and MacFarlane (2010) famously describe the following case. Ten miners are trapped in a shaft—*A* or *B*, we don't know which—and threatened by rising waters. We have time to block exactly one shaft or neither. If we block the correct shaft, everyone lives. If we block the wrong shaft, everyone dies. If we do nothing, exactly one miner dies. Thus:

	in_A	*in_B*
block_A	All live	All die
block_B	All die	All live
$\neg$(*block_A* $\vee$ *block_B*)	Nine live	Nine live

This is actually an elaboration of the sort of case described in the previous section. Your first-order preferences among outcomes look like this:

$(in\_x \wedge block\_x) \prec_{FO} \neg(block\_A \vee block\_B) \prec_{FO} (in\_x \wedge block\_y)$

Now, *what should we do*? Absent information about the probability, e.g., that the miners are in A, what we should do is *indeterminate* (or, perhaps, it is determinate that we should block neither). But evidently the best possibilities are all possibilities where all miners live. So the best possibilities are all possibilities where we block A or block B. Apparently, then, according to the Standard Account, the following is true:

(6) We should either block A or block B. $\Box(block\_A \vee block\_B)$

Not only does this sound false (or at least not determinately true); it seems that, on certain assignments of probabilities to the claim that the miners are in A (e.g., 50%), the following is instead true:

(7) We should do nothing. $\Box\neg(block\_A \vee block\_B)$

As we discussed above, a fan of the Standard Account might reply: there are things, other than the lives of the miners, but which we manifestly do care about (e.g., dealing with uncertainty, managing risk), that the decision table fails to represent. In this vein, von Fintel (2012), addressing cases specifically in which we judge (7) true, writes: "It is not obvious that [a world where all ten miners are saved] is a better world than one where nine miners survive because we choose not to run the risk." Thus our considered preferences are, in fact, of the following shape:

$$\neg(block\_A \vee block\_B) \prec_{ATC} (block\_A \vee block\_B)$$

This, the claim goes, underwrites our judgment of the truth of (7) and the falsity of (6).

Impressively, this sort of thought can be extended to explain why we tend ("instantly," in the phrase of von Fintel (2012)) to judge the following conditionals true:

(8) If the miners are in A, we should block A. $(if\ in\_A)(\Box block\_A)$

(9) If the miners are in B, we should block B. $(if\ in\_B)(\Box block\_B)$

All we need do is combine it with the further thought—for which there is an ever-expanding body of evidence (see, e.g., Yalcin 2007; Gillies 2010; Charlow 2013a, b)—that the *if*-clause of an indicative conditional functions, in part, to shift the information relevant for evaluating information-sensitive expressions (such as DNMs). This allows us to exploit this claim of Kolodny and MacFarlane (2010):

> A world may be more ideal than another relative to one information state and less ideal than it relative to another. For example, a world in which both shafts are left open may be more ideal than one in which shaft A is closed relative to a less informed state, but less ideal relative to a more informed state. (p. 133)

Let $\prec_{ATC[\phi]}$ give our considered preferences given a hypothetical update with information ϕ—i.e., the sort of update induced by the *if*-clause of an indicative conditional

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 59 — #13

of the form $(if\,\phi)(\psi)$. If we take on Kolodny and MacFarlane's quoted claim, we have a rationale for endorsing the following:

$$block\_x \prec_{ATC[in\_x]} \neg(block\_A \vee block\_B) \prec_{ATC[in\_x]} block\_y$$

Bracketing details, it isn't hard to use this sort of claim about the structure of the suppositional or hypothetical preference order to underwrite a semantics for conditional DNMs on which (8) and (9) come out true, even while (6) comes out false (relevant formal tools can be found in Charlow 2013b; Cariani et al. 2013; Silk 2014).[13]

4.4 Decision theory and truth-conditions

Here we have a Quietist account of the Miners Puzzle that seems, at a first pass, adequate. Kratzer's question presents itself with even greater force: "Why pack information about rational decision making into the meaning of modals?" Here is a response (also pursued in Charlow 2013b).

A preliminary answer: in fact, it seems that a key motive for introducing orderings into the semantics for DNMs is to *recapitulate*, if partially, the structure of practical reasoning. Practical reasoning takes place against preferences over outcomes as well as information about outcome-relevant circumstances. This explains, in part, why the semantics of DNMs must (and does) distinguish between something analogous to a preference-state and something analogous to an information-state. Note, further, the fundamental theoretical purpose of a normative **decision theory** to bridge the gap between, on the one hand, descriptions of the agent's preferences over outcomes (and the agent's information about outcome-relevant circumstances) and, on the other hand, an account of what that agent should decide, intend, do. (Important note to the reader: by "decision theory" I will mean nothing more than *something that determines a choice function*. A choice function is an object that, given a specification of preferences over outcomes and information about outcome-relevant circumstances as input, generates a set of permissible actions as its output.)

If the recapitulation claim is right, there is reason for the semantics of DNMs—claims which serve to express verdicts about what agents should do—to represent the normative role of a decision theory. Note that, in a non-semantic context, it would be something like a category error to build the decision theorist's conception of the preferred way to bridge the gap between preference/information and action into the description of what outcomes are preferable for the agent. Further, it is clear that verdicts of what an agent *should do*, in light of a description of what outcomes

[13] Kolodny and MacFarlane (2010) argue that this means giving up modus ponens, since $(in\_A \vee in\_B)$ together with (8) and (9) apparently entail (6). But this view—note I include my concurrence in Charlow (2013b)—is mistaken. As Willer (2012) notes, if our notion of entailment is dynamic in nature—roughly, $\phi_1, \ldots, \phi_n \models_{Dyn} \psi$ iff the result of updating an information-state on $\phi_1, \ldots, \phi_n$ yields an information-state that accepts ψ—this does not follow. Given one very natural treatment of the conditional, on which a state σ accepts $(if\,\phi)(\psi)$ iff the result of updating σ with ϕ accepts ψ—modus ponens is dynamically valid—$(if\,\phi)(\psi), \phi \models_{Dyn} \psi$. But constructive dilemma is clearly *not*—$(if\,\phi)(\chi), (if\,\psi)(\rho), (\phi \vee \psi) \not\models_{Dyn} (\chi \vee \rho)$. For some discussion, see Yalcin (2012b, n. 14); Charlow (2013a, §4.2).

are preferable for the agent (and the information to which she has access), must be generated by appeal to some decision theory or other. What goes for decision theory goes, then, for the theory of what the agent should do: take care to distinguish between first-order and rational preferences. There is, I would think, strong pressure to say that what goes for the theory of what the agent should do goes, in turn, for a semantic theory for sentences that conventionally serve to express views about the verdicts of the theory of what the agent should do. Regardless, then, of whether we are doing decision theory or theorizing about the meaning of modals that express the considered verdicts of a decision theory, it seems we should take care to avoid building rational preferences into our description of first-order preferences. But it seems like this is exactly the conflation Kratzer is encouraging us to make.

To put it differently, consider a decision problem Π for an agent with preferences P_Π, information I_Π (which also specifies a range of available actions for the agent, $A_\Pi = X_1, ..., X_n$); we will represent the decision problem as an ordered triple of the form $\Pi = \langle P_\Pi, I_\Pi, A_\Pi \rangle$. Let D be the selected decision theory,[14] and let $\mathcal{CH}_D$ be the choice function determined by D from decision problems into sets of D-admissible actions in a decision problem. Now consider:

The Master Argument for "Packing In" Decision Theory

i. $\Diamond X$ (given Π) iff $X \in \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$
ii. "$\Box X$" is true at Π iff $\Box X$ (given Π)
iii. $\Box X$ (given Π) iff $\neg\Diamond\neg X$ (given Π)
iv. So "$\Box X$" is true at Π iff $\neg X \notin \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$

Because which actions are permissible within a decision problem is in part a function of a choice of decision theory (i), by an application of the T-Schema[15] (ii) and the metalinguistic duality of deontic necessity and possibility (iii), the truth-condition of a sentence describing what an agent should do in a decision problem is also such a function (iv).[16]

[14] D can be any decision theory whatever—any recipe from transforming preferences over outcomes and information into a choice function, including recipes that, e.g., treat information as irrelevant. D need not be an expected utility calculus, or indeed something borrowed from the decision-theoretic literature at all.

[15] Instances of (ii) are not, strictly, instances of the T-Schema, which, as Fabrizio Cariani has pointed out to me, concerns the application of a monadic, not relational, truth-predicate. Still: schema (ii) is correct and clearly related to the T-Schema (and seems to me to be endorsed by Tarski himself when he discusses monadic *truth* as a special case of relational *satisfaction*). Consider an index-sensitive truth-condition for the sentence *grass is green*. This sentence is true at w iff $[\lambda v.$*grass is green at* $v](w) = 1$ iff grass is green at w. Generalizing, $\ulcorner\phi\urcorner$ is true at some sequence of indices S iff $[\lambda S'.\phi(S')](S) = 1$ iff $\phi(S)$.

[16] I concede, of course, that work needs to be done to fill out this basic picture. Here I will mention two outstanding issues. First, if it turns out that which things are chairs depends on the distribution of a certain kind of fundamental physical particle P, this argument could thereby be used to show the need to represent P in the semantics of a sentence such as *that's a chair*. That would be a reductio of this argument. (I regret that I cannot recall the name of the person who raised this worry to me.) Avoiding this sort of overgeneration is not

Two related lessons can be drawn from this. First, if a kind of parameter—more precisely, a kind of distinction between parameters—is constitutively relevant to the determination of what the agent should do, it is *thereby semantically relevant* to the truth-condition of a sentence describing what she should do. Second, a semantics that refuses to distinguish between first-order and rational preferences cannot represent the distinct contribution of first-order and rational preferences to the truth-condition of a DNM—a state of affairs that is at once both theoretically embarrassing and liable to lead to empirical difficulties. (I will outline an empirical difficulty below.)

This is the case for "pack[ing] information about rational decision making" into the meanings of modals—at least into the meanings of modals that express verdicts about rational decision making. Quoting myself, cases like the Miners (and the schematic case on which it is based) . . .

> . . . recommend the introduction of some sort of *independent theoretical apparatus* taking some specification of (i) substantive, first-order ends and (ii) an information-state as its input, and generating a modified ordering on possibilities as output. We could, of course, conflate what is "best" or "ideal" with what is minimal with respect to the modified ordering [i.e. $\prec_{ATC}$]. Semantically, this would amount to restyling the deontic selection [*min*] as selecting, not worlds that are best sans phrase [i.e. relative to $\prec_{FO}$], but with respect to some mixture of [first-order and rational preferences] . . . [But we] should take care to avoid running together theoretical notions (and pieces of theoretical apparatus) that, like actual and expected utility, a theory ought not to run together. (Charlow 2013b, pp. 2309ff.)

To be clear, I am not arguing against allowing $\prec_{ATC}$, rather than $\prec_{FO}$, to determine the DNM's domain of quantification. I am arguing against a decision-theoretically quietest semantics—one that refuses to represent anything more than $\prec_{ATC}$ in the truth-condition for the DNM. Even if we allow $\prec_{ATC}$ to determine the domain of quantification for the DNM, *we must also represent the distinct contributions of (at least) first-order preferences and rational preferences in determining* $\prec_{ATC}$. A story that fails to represent the distinctive roles of first-order preferences and preferences concerning ways of reasoning about what to do in the truth-conditions of claims invoking DNMs cannot be the whole story about their meanings.

so hard, if we understand the premises of the argument as involving special equivalencies—e.g., conceptual or analytic—that do not hold between chairs (and claims about chairs) and claims about *P*s.

Second, $\mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$ must be closed under conditions that are appropriate given the logic of DMs. For instance, if $\Diamond X$ entails $\Diamond Y$, then for any $\mathcal{CH}_D$ and Π such that $X \in \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$, it must be that $Y \in \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$. Giving a precise description of the operations under which $\mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$ is closed requires settling quite a lot of disputed issues in the logic of DMs—e.g., whether permission itself is closed under arbitrary disjunction (i.e., the problem of Free Choice Permission)—and I will avoid it here. I will note that imposing closure conditions on $\mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$ will make it less natural to say that the elements of $\mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$ are *actions*—rather, they will be *constructs out of* the range of permissible actions delivered directly by the relevant decision theory, often, but not always, corresponding themselves to actions. In no way does this vitiate the argument of this section: a DM expresses a judgment about what should be done, which in turn depends on which actions are permissible in light of the relevant decision theory.

4.5 Rational fetishists

To see this better, I will describe an empirical difficulty. Note first that von Fintel's claim that it's "not obvious that [a world where all ten miners are saved] is a better world than one where nine miners survive because we choose not to run the risk" (and Kolodny and MacFarlane's parallel claim that "a world in which both shafts are left open may be more ideal than one in which shaft *A* is closed relative to a less informed state") must be wrong. In the most immediate senses of *better* and *ideal*, the best, or most ideal, outcome is the outcome in which all ten miners live; the central difficulty of the case is that we don't know how to bring the best outcome about.

Not, I think, that anyone would deny this. But it is easy to imagine someone—call him the *Rational Fetishist*—who did. The Fetishist values ordinary things (for instance human life), as well as the demands of practical rationality, in *exactly the same way*. For the Fetishist, ordinary things do not function as constraints on the pursuit of the demands of practical rationality; these kinds of goals are comparable to (and, so, on occasion, compete with) one another. (For the ordinary rational agent, the demands of practical rationality do not compete with non-decision-theoretic goals; the former rather structures the agent's rational pursuit of the latter.) For the Fetishist, the value of not risking the deaths of all ten miners in fact renders the outcome in which we block neither shaft actually better, literally more desirable, than the outcome in which we block the correct shaft.

Note that, so long as the same decision theory (hence the same decision-theoretic verdicts) applies to the ordinary rational agent and the Fetishist, the considered preferences of the Fetishist coincide—indeed, if it helps, we could imagine that they necessarily coincide—with the considered preferences of the ordinary agent. But I submit the following as a theoretical datum (although I will also argue for it): *the truth-condition of a DNM "in the mouth of" the Fetishist is distinct from the truth-condition of a DNM "in the mouth of" the ordinary agent*.[17] A theory which refuses to assign distinctive roles to first-order preferences and rational preferences in the truth-conditions of claims invoking DNMs will find this hard to explain (since in such a theory only $\prec_{ATC}$ is visible, and the properties of this ordering *coincide* for the kinds of agents under consideration).

In support of this datum, note that, for the Rational Fetishist, there is, by stipulation, a *single* preference-structure in play, whose implementation in, e.g., the Miners scenario, requires blocking neither shaft. It is a necessary condition on implementing the relevant preference-structure that neither shaft be blocked. According to at least

[17] To be clear: I do not really mean to assume that the preference-structure relevant for interpreting a DNM is supplied by the agent who *utters* the DNM (hence the scare quotes around "in the mouth of"). Here I intend something much weaker: that when we (qua semantic theorists) evaluate DNMs relative to orderings that are supplied by Fetishistic and ordinary preference-structures, we should be able to identify a semantic distinction between them.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 63 — #17

one influential story about the semantics of strong DNMs (e.g., *must*) (see von Fintel and Iatridou 2008),[18] a claim of roughly the form in (10). . .

(10) To implement the basic preferences, we **must** block neither shaft.

. . .is therefore true relative to a Rationally Fetishistic preference-structure. For the ordinary agent, however, there are *two* preference-structures in play: first-order and decision-theoretic (the former valuing things like miners' lives, the latter perhaps valuing things like risk-sensitivity in pursuit of the former). It is *not* a necessary condition on implementing the "primary" preference-structure that neither shaft be blocked (in fact, as we've seen, it is a necessary condition on implementing the "primary" preference-structure that at least one shaft be blocked). Blocking neither shaft is, however, the best way of implementing the primary preference-structure, given risk-averse rational preferences. So, applying the von Fintel and Iatridou (2008) account of the semantics of weak DNMs (e.g. *should*), a claim of roughly the form in (11) is therefore true when evaluated relative to an ordinary preference-structure.

(11) To implement the basic preferences, we **should** block neither shaft.

However, a claim of the form in (10) is judged (appropriately, in my view) false.[19]

There is, then, a fairly clear semantic difference between DNMs "in the mouths of" Rational Fetishists and those "in the months of" ordinary agents. To explain this, it seems we should augment the Standard Semantics with some of the fruits of normative decision theory. For instance, in Charlow (2013b), to account for our ordinary judgments about the Miners scenario, I suggested making use of a truth-condition inspired by the truth-condition in Q3, repeated here:

(Q3) $$min_{\preceq_2}(min_{\preceq_1}(R)) \subseteq [\![p]\!]$$

More precisely, I suggested semantically representing pairs of orderings—and, potentially, sequences of orderings of arbitrary length—and suggested a recipe for constructing a *mixed ordering*—corresponding here to the considered preference ordering—from the sequence together with some specification of the relevant information. (In this treatment, the mixed ordering was information-sensitive in roughly the way described earlier in this section.) My aim was to state a truth-condition for the DNM that tracked the verdict of a specific kind of decision theory—the knowledge-based decision-theory of Hawthorne and Stanley (2008)—in the case. But other views,

[18] Note: none of what I say here requires taking on the treatment of the strong/weak distinction in von Fintel and Iatridou (2008), to which I will object below. My view is that on *any* adequate treatment of the strong/weak distinction, (10) should come out true relative to a Rationally Fetishistic preference-structure, false relative to an ordinary preference-structure. I will not attempt to defend any view of the strong/weak distinction here.

[19] For more on the importance of the strong/weak distinction in the Miners scenario, see Charlow (2013b); Silk (2014).

making use of different decision theories, could easily be plugged into this basic apparatus.

Signpost. I have argued that a semantics for DNMs should have, in the first instance, the following sort of shape: $\Box X$ is true (given Π according to D) just if $\neg X$ is impermissible in Π by lights of D:

$\Box X$ is true (given Π) iff $\neg X \notin \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$

The truth-condition of a DNM will reference a choice function $\mathcal{CH}_D$, supplied by the relevant decision theory D, and a decision problem $\Pi = \langle P_\Pi, I_\Pi, A_\Pi \rangle$, given presumably at a context.

Before moving on, I wish to emphasize, briefly, that this in no sense involves a departure from the suggested lesson of §3. Rather, for all we have seen, it reinforces it. The Standard Semantics is too coarse-grained to represent phenomena of normative significance which seem to bear clearly on the truth-conditions of DNMs. The way to fix this is to refine it—in particular, by augmenting it with the fruits of decision theory.[20]

Nor does it necessarily involve any departure from the idea that DNMs have *quantificational truth-conditions* (cf. §2). We are free to pursue an equivalence between a quantificational truth-condition and a decision-theoretic truth-condition of the form given here (analogous to the equivalence demonstrated in §2).[21] (While I am doubtful we can retain anything resembling Kratzer's treatment of comparative betterness, as encoded in Definitions 4 and 5, the assumptions that endow these definitions with predictive power are—as I explain in §6.4; see also Lassiter (2011)—independently problematic. It is, in my view, a virtue of the account I am suggesting here that we can move beyond these definitions. The account I advocate achieves predictiveness without taking on problematic stipulations about the properties of the relevant ordering.)

Here we have what I take to be an intuitive, well-supported, and altogether fairly conservative picture of how theorizing about the semantics of DNMs should proceed. There are, however, deep difficulties in implementing it. I describe these in the next section.

[20] As I emphasized in Charlow (2013b), the representation of decision theories in the semantics of weak DNMs is motivated by the same sorts of considerations that motivate the von Fintel and Iatridou analysis of weak DNMs (§3.1). For this reason it is puzzling that von Fintel (2012) portrays the kind of analysis advocated in this section as a "challenge" to the classical semantics for DNMs (since he does not portray his own understanding of weak DNMs similarly). The view I have defended does rely on a challenge to the classical (Kratzer) semantics for indicative conditionals, but these are very different things.

[21] How would that work? A choice function $\mathcal{CH}_D$, when applied to a decision problem Π, selects a set of alternatives that are permissible according to D. The "disjunction" of these alternatives $\bigcup \mathcal{CH}_D(\Pi)$ represents something like the "sphere of permissibility" (compare Lewis 1979b) for the agent in Π, which is formally akin to an ordinary domain of quantification, and which can therefore be linked to claims of comparative betterness using the sort of strategy pursued in §2.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 65 — #19

5 Packing It In

The difficulties begin with the observation that this schematic sort of truth-condition cannot be, strictly speaking, correct. It ignores a range of data suggesting that DNMs have context- and index-sensitive truth-conditions. Here I will connect this rather prosaic fact about DNMs to Carr's (2012a, b) (in my view successful) objection that all semantics for DNMs in the mold I have described objectionably write normative assumptions into the semantics.[22] The scope of Carr's objection is, I shall argue, rather greater than you might expect. I shall argue for this by attempting to describe a proposal broadly along the lines she suggests (on which DNMs are parametrized, not only to decision problems, but to decision theories). It is fairly clear that taking this sort of proposal seriously will mean excising *all* "substantive" normative assumptions—basically, I will argue, all the normative assumptions—from the semantics. This is a delicate matter, since it is, I'll argue, prima facie incompatible with the sort of research program sketched in the prior sections.

5.1 Disagreement

It is a truism that agents with different preferences and priorities will tend toward disagreement about what is to be done. It is also a truism that such agents are generally linguistically competent with respect to DNMs and the broadly normative concepts they function to express (note: I will henceforth refer to this mixture of linguistic and conceptual competence simply as *competence* with respect to DNMs). While it would be too strong to say that it is the job of the *semantic* theorist to *account for* this sort of disagreement—here I disagree with, e.g., MacFarlane (2011) for reasons broadly like those given by Plunkett and Sundell (2013)—it is at least incumbent on the semantic theorist not to state a theory which renders the disagreement unintelligible (compare Gibbard 1990, chapter 1).

Consider the Newcomb Problem. For the uninitiated, a brief description. There are two boxes, *A* and *B*. You get to keep whatever is inside any box you open. There is a powerful **Predictor**, which makes extremely reliable predictions about your actions, which at the time of your decision has already acted thus:

i. It has put \$1, 000 in box *A*.
ii. If it has predicted you will open just *B*, it has in addition put \$1,000,000 in *B*.
iii. If it has predicted you will open both boxes, it has put nothing in box *B*.

[22] I take myself here to be fleshing out an objection belonging to Carr (2012a, b); this is also part of the project of Cariani (2016), as I note at many points below. Carr's main worry about writing a normative assumption *N* into the semantics of DNMs is that it predicts a kind of linguistic incompetence in those who use DNMs to express normative views distinct from *N*. (Intriguingly, she connects this to G. E. Moore's Open Question Argument against naturalistic analyses of "good".) I quite agree with her. I only claim that there is a lot worth saying about the issue beyond what Carr has said about it. I will try to say a small portion of it here. I've no reason for thinking she would disagree with any of it.

At the time of your decision, you know the Predictor has already acted thus. You are given a choice between two actions: opening both boxes (two-boxing), or opening just box *B* (one-boxing). The decision problem is given by roughly the following table:

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

Following Gibbard and Harper (1981), Evidential Decision-Theorists ("EDT-ers") calculate the expected utility of an action X using, for each contingency C, the probability of C conditional on the agent doing X, and they tend to assert (12). Causal Decision-Theorists ("CDT-ers") calculate the expected utility of an action X using, for each contingency C, the probability of C were the agent to do X—which, when the agent's action does not causally affect the probability of C, as in the Newcomb Problem, is simply equal to the probability of C. CDT-ers tend to assert (13):

(12) You should one-box.

(13) You should two-box.

Truism 1: each speaks correctly (in the minimal sense that each appropriately deploys her competence to express her view about what to do). Truism 2: they disagree—about what to do. Final truism: the disagreement stems from different views about which decision theory—EDT or CDT—is correct. Assuming the truth-condition of the DNM is of the shape for which I have argued, the disagreement originates, not from a disagreement about what value to plug in for P_{Π}—all parties agree about the shape of the decision problem—but rather from disagreement about what value to plug in for $\mathcal{CH}_D$.

Similar phenomena will, I conjecture, arise for any normative view about which there can exist disagreement among agents that are competent with respect to DNMs. (Note that whether the disagreement is rooted in one party's irrationality is neither here nor there, so long as the disagreement *exists* in the presence of the relevant semantic and conceptual competence.) This plausibly describes any contentful normative view whatever.

5.2 Binding

In what will turn out to be a similar vein, next consider (cf. Carr 2012a, p. 181):

(14) If EDT is correct, you should one-box.

(15) If CDT is correct, you should two-box.

These claims are intelligible and seem to have an obvious meaning. (In fact, they are apparently *correct* to say.) To them we might add:

(16) If the correct decision-theory is risk-averse, we should block neither shaft.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 67 — #21

(17) If the correct decision-theory is risk-seeking, we should block a shaft.

Similar conditionals can be formulated and used to intelligibly (and correctly) express claims about the verdicts of any contentful normative view whatever for any decision problem whatever.

Such conditionals would, on the face of it, seem likely to be compositionally similar to more prosaic examples such as:

(18) If the law forbids jaywalking, you should wait for the light.

(19) If you want to go to Harlem, you should take the A-Train.

These examples are apparently united by this feature: a hypothetically introduced decision theory—in cases (18) and (19), a hypothetically introduced preference—apparently "binds" (in a very loose sense) a corresponding parameter in the matrix DNM.[23]

5.3 Parametrization

What morals should a semanticist draw from such cases? While Carr (2012a, b) perhaps does not have in mind this exact question, her admonition against writing substantive decision-theoretic assumptions into the semantics (and suggestion to parametrize the truth-conditions of DNMs to decision theories in addition to decision problems) gives clear guidance. We cannot represent the disagreement in (12) and (13) as intelligible, nor can we understand the apparent ability of hypothetically introduced preferences to bind corresponding parameters in the matrix DNM, unless we allow—contra a large body of orthodoxy—that the semantic variability in a DNM—a vague notion, but one we will precisify shortly—outruns the specification of the relevant decision problem.

There are various ways to see this. Let Π be the Newcomb Problem, and let $\mathcal{CH}_{\mathrm{EDT}}$ and $\mathcal{CH}_{\mathrm{CDT}}$ be the choice functions of EDT and CDT respectively. Suppose further that the *correct* decision theory is CDT, and suppose we take this to motivate a truth-condition for DNMs of this shape:

$\Box X$ is true (given Π) iff $\neg X \notin \mathcal{CH}_{\mathrm{CDT}}(P_{\Pi}, I_{\Pi}, A_{\Pi})$

If CDT is the correct decision theory, this truth-condition is, in a very natural sense, the *correct truth-condition* for the DNM. (I'll elaborate below.) But it would seem that we must resist this very natural thought. It seems to prevent us from understanding the EDT-er's utterance in (12). The truth-condition the EDT-er intends is apparently

[23] For differing views of how this "binding" happens, see von Fintel and Iatridou (2005) and Charlow (2013a). In the latter, I argue that the sort of doubly modalized account of these sorts of conditionals sketched in von Fintel and Iatridou (2005) cannot give a fully general account of the semantics of these conditionals, roughly because it does not take the notion that this is a form of binding seriously enough. In a similar vein, Yalcin (2012a); Cariani (2016) cite the ability of probabilistic claims in antecedents to affect the interpretation of modals in their consequents to motivate a treatment of probabilistic semantic parameters akin to the treatment of the decision-theoretic semantic parameter for which I argue here.

semantically ineligible—ruled out as a potential meaning of the DNM, by semantic fiat. Insofar as one party to the disagreement cannot be interpreted as meaning something intelligible, this apparently commits us to regarding the disagreement between (12) and (13) as unintelligible. This is a disaster. People can, of course, express mistaken normative views and engage in substantive normative disagreements by using (and thereby exploiting their competence with respect to) DNMs.[24]

How can we accommodate decision-theoretic variability in the semantics? We have options. Here, first, is a non-starter. Suppose we intensionalize $\mathcal{CH}$, so that it expresses a function from an ordinary circumstance of evaluation i (e.g., a world or world-time pair) into the the choice function characterized by *the correct decision theory* at i.

$\Box X$ is true at i (given Π) iff $\neg X \notin \mathcal{CH}(i)(P_\Pi, I_\Pi, A_\Pi)$

This gets certain things right. It allows, for example, that the EDT-er manages to express a proposition. More impressively, if we simply assume a vanilla strict-conditional semantics for indicative conditionals—so that $(\textit{if}\, p)(q)$ is true iff, roughly, for all i such that p is true at i, q is also true at i—intensionalization renders the truth-condition of a conditional such as (14) roughly as follows: for all i such that EDT is the correct decision theory at i, the DNM *you should one-box* is true at i. This is not bad; indeed, it is quite similar to the truth-condition suggested by, e.g., von Fintel and Iatridou (2005) for preference-shifting, "anankastic" conditionals such as (19).

But it gets certain things badly wrong. The proposition expressed by the EDT-er who asserts (12) is apparently not one she would intend to express: it is a proposition true at i just if $\mathcal{CH}(i)$ (which is by assumption the choice function of CDT for arbitrary i) selects one-boxing. It is a proposition that is, moreover, false with respect to any circumstance of evaluation (since, as everyone knows, CDT tells you unequivocally to two-box). Why, then, would an EDT-er care to assert it? (Similar remarks apply mutatis mutandis to a theory which treats $\mathcal{CH}$ as expressing a function from a context of utterance c into the correct decision theory at c.) It is, in any case, not the right story about the sort of binding into the matrix DNM involved in these sorts of constructions more generally (for further details, see Charlow 2013a).

A more promising strategy—and the one Carr herself adopts—simply introduces a decision theory as a *separate and sui generis semantic parameter*.[25] (I'll refer to this

[24] Similarly, Cariani (2016) argues against a "Fully Bayesian" attempt to write a decision theory into the semantics on DNMs (on which, roughly, $\Box X$ says the expected value of X is sufficiently high), by noting that a Fully Bayesian semantics for DNMs cannot intelligibly represent the disagreement between someone whose choices are characterized by a risk-averse choice function and someone whose choices are characterized by a Fully Bayesian choice function. (Cariani additionally offers a convincing response to any attempt to represent the view of the risk-averse agent as Fully Bayesian.)

[25] The positive proposal of Cariani (2016) is an interesting direct challenge to this claim, since it does not involve the use of any decision-theory parameter, but still claims to achieve the right kind of decision-theoretic neutrality. Apart from the sort of neutrality Cariani is interested in, however—which is in his essay limited to being able to represent attitudes of and disagreements between agents with different attitudes toward risk—there is still the issue of accounting for the meaning of (12)–(15). On this matter: I think there is likely a tension between Cariani's use of a binding-type argument to motivate parametrization to probability

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 69 — #23

strategy as Parametrization or Relativization.) There are different ways this sort of strategy could be pursued. According to Dowell-style Contextualists (Dowell 2011), we would parametrize the *content* of a DNM, not just to decision problems, but also to decision theories. MacFarlane-style Relativists (e.g. Stephenson 2007; MacFarlane 2011) will also want to parametrize DNMs to decision theories, but will construe this further parametrization as an elaboration of the context in which DNMs are *assessed* for truth (rather than the context in which they are assigned interpretations). Various other positions on the nature and proper theoretical interpretation of this parametrization (and the mechanism by which conditionally introduced preferences bind parameters in matrix DNMs) are possible.

5.4 Competing truth-conditions

Parametrization to decision theories seems to me the way to proceed here. But there is a rather fundamental difficulty with it. Parametrization to decision theories would seem to entail allowing the possibility of a *competing truth-condition* associated with a DNM "in the mouth of" the EDT-er, of roughly the following shape:

$\Box X$ is true (given Π) iff $\neg X \notin \mathcal{CH}_{\mathrm{EDT}}(P_{\Pi}, I_{\Pi}, A_{\Pi})$

Prima facie, it thus seems to mean allowing that the EDT-er speaks *truly* with (12). Of course that is something *we* should feel uncomfortable saying, if CDT is the correct decision theory. If CDT is the correct decision theory, (12)—even "in the mouth of" the EDT-er—is simply false.[26] Indeed, this is precisely what the conditional (15) seems to be (correctly) expressing. A similar sort of problem will arise in any case in which conditionals can be formulated and used to express claims about the verdicts of a contentful normative view.

Here I want to work through some options for dealing with this general sort of problem (though I will maintain my focus on the issues arising from the EDT/CDT disagreement). I will mostly be re-canvassing territory already covered in the Contextualism-Relativism literature.

functions (see fn. 23) and his decision not to parametrize to choice functions (since similar arguments can be mustered for both).

[26] One may object to the apparent assumption that the decision theory "of" the speaker determines the decision-theoretic parameter against which the speaker's assertion of (12) is evaluated for truth. But there is still the problem that evaluating what the EDT-er asserts, relative to $\mathcal{CH}_{\mathrm{EDT}}$ yields a verdict of truth. To avoid the conclusion that the EDT-er speaks truly, at, we need to *prohibit* evaluation of what the EDT-er asserts relative to $\mathcal{CH}_{\mathrm{EDT}}$. It is unclear on what grounds such a prohibition might rest. Phrased this way, I suspect this is simply a vivid instance of a more general difficulty with model-theoretic semantics—its difficulty deriving truth-conditions for false sentences *on which they can be seen to come out false*. As Lepore (1983, p. 178) puts it, "[F]rom a relativized truth-theory we cannot derive an absolute truth-theory." I do not want to be construed as endorsing this critique generally, but I think it is powerful in the case of DNMs.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 70 — #24

No correct view? Perhaps there is no fact of the matter about what the correct decision theory is (perhaps owing to a meta-normative Error Theory, à la Mackie 1977).[27] If this is right, then the supposition that CDT is correct is in some sense an ineligible hypothetical.

This would come as a surprise to partisans of the CDT/EDT debate. It seems to have the further consequence that commitments to and assertions of DNMs that are, on the face of it, premised on an assumption of the correctness of either CDT or EDT—(12) and (13)—cannot be rationally maintained. I take this provisionally as a reductio of the view under consideration. Whatever the meta-normative facts turn out to be, the possibility of rational commitment to the correctness of a normative view (and the ability to rationally assert this commitment) should not be relinquished.[28] At the least, the notion that recherché metasemantic considerations such as those in play here would prompt us to relinquish it should not, absent much further argument, be taken seriously.

Ecumenism? Perhaps, despite appearances, both EDT-ers and CDT-ers speak correctly (by which I mean something more than that they appropriately deploy their competence with respect to DNMs). There are two main ways of making good on this thought: Contextualism and Relativism.[29]

According to Contextualism, the EDT-er who asserts (12) makes a claim whose content is glossed roughly as *the choice function of EDT selects one-boxing*, while the CDT-er who asserts (13) makes a claim whose content is glossed roughly as *the choice function of CDT selects two-boxing*. But, while I acknowledge that there are theorists who would disagree, I will suppose that this does not correctly report the content of the CDT-er's claim. This would render the content of the CDT-er's assertion of (13) *trivial* (likewise for the EDT-er's assertion). It would thus render the purpose of any article in which a CDT-er undertook to defend the correctness of such an assertion

[27] This sort of response works better in cases involving disagreements with epistemically modal claims. There is no fact of the matter about what the *correct* domain of quantification for epistemic modals is: there is nothing over and above different bodies of evidence from whose vantage different epistemically modal claims can be made. (A relatively more informative body of evidence is in some sense superior to a relatively less informative body of evidence, but taking this idea seriously in a metasemantics for epistemic modals will ultimately require us to shrink domains of quantification to maximally informative domains, i.e., singleton sets of possibilities.) Similar remarks could be made for other sources of variability in epistemic claims. For instance, assuming that knowledge-attributions are sensitive to stakes or to standards, most epistemologists would agree that it is nonsensical to ask *which stakes or standards* are correct. The most we can say as theorists is to describe which certain stakes or standards are operative in a context, by way of explaining the judgments of truth and falsity that speakers who are competent with the relevant epistemic language tend to make. (The skeptic, of course, disagrees with this general assessment, but this has earned the skeptic a fair share of opprobrium in the contemporary epistemological literature on skepticism.) For reasons that will emerge shortly, this is also why I think Contextualism about epistemic language remains viable in the face of this sort of worry, while Contextualism about deontic modals does not.

[28] For a defense of such commitments (within a broader program of anti-Realism about the normative) that I take to be basically sound, see Blackburn (1984, 1998); Gibbard (1990, 2003).

[29] There are, of course, many more than two ways, but I must confine my attention here to the main contenders.

mysterious. This is not exactly a palatable way to regard the work of your philosophical colleagues.[30]

According to Relativism, however, the EDT-er who asserts (12) makes a claim whose propositional content is glossed roughly as *the correct choice function selects one-boxing*.[31] This seems like exactly the right thing to say. For the Relativist, the truth of this content is, however, said to be perspective-dependent. In particular, it is dependent on a choice of choice function. The claim *the correct choice function selects one-boxing* is true relative to $\mathcal{CH}_{\mathrm{EDT}}$, false relative to $\mathcal{CH}_{\mathrm{CDT}}$. Thus both the EDT-er and the CDT-er speak correctly, since both assert contents that are true relative to the choice functions supplied by their respective perspectives. So Relativism is a form of Ecumenism.

Relativist Ecumenism seems to offer the best hope of a propositional semantics for DNMs that is responsive to the worry developed in this section. In the next section, I will explain why it is not a very good hope and go on to explain the implications of this for theorizing about DNMs.

6 Metasemantics

In this last section, I will first explain my dissatisfaction with propositional semantics for DNMs. I will suggest a semantics—really, a metasemantics—that, rather than attempting to deliver "truth-conditions" for DNMs, simply attempts to *model linguistic competence* with respect to DNMs: to represent (very roughly) what state of mind an agent is in when she judges a DNM to be the case. Because such competence can co-exist with effectively any normative view, the core semantics for DNMs must be extremely minimal: the compositional semantics of a DNM of the form $\Box\phi$ is nothing more than a function from sequences of parameters $P_1, \ldots, P_n$ into a

[30] One suspects Plunkett and Sundell (2013) will deploy an alternative way of rationalizing the assertion—e.g., as an attempt to achieve a mutually acceptable conversational scoreboard (compare Lewis 1979c) on which $\mathcal{CH}_{\mathrm{CDT}}$ is mutually agreed to partly determine the content of assertions of DNMs, or as an attempt to persuade one's interlocutors to modify their own decision-theoretic views. (Their preferred term here is "negotiation.") I am, as a rule, basically happy with these sorts of pragmatic stories, but, for reasons I will describe below, I think it makes better sense to embed them in an Expressivist (Charlow 2014b, 2015; Yalcin 2012a; Rothschild 2012) or Constraint-Theoretic (Swanson forthcoming) metasemantics.

I note that the most convincing defenders of Contextualism for epistemic modals—von Fintel and Gillies—choose as their *ur*-semantics for epistemic modals an apparently non-truth-conditional dynamic-semantics borrowed from Veltman (1996) (see, e.g., von Fintel and Gillies 2007). Insofar as Contextualism about epistemic modals is committed to the doctrine that epistemic modals conventionally express propositions, their view is simply not a Contextualist view. However, insofar as Contextualism about epistemic modals is *not* committed to the doctrine that epistemic modals conventionally express propositions, with all of the dubious pragmatic commitments that are normally thought to come along with such a commitment—e.g., that the essential effect of accepting an epistemic modal involves adding the proposition it expresses to one's beliefs (against this, see esp. Yalcin 2007)—I would not object to being called a Contextualist.

[31] As this suggests, I am understanding Relativism to encompass both MacFarlane-style Assessor Relativism and various forms of Nonindexical Contextualism. The differences between these views do not, I think, matter for my purposes here.

condition glossable, roughly, as *the domain characterized by* $P_1, \ldots, P_n$ *is* ϕ*-entailing* (equivalently: $\neg\phi$ is not among the set of admissibles characterized by $P_1, \ldots, P_n$). Fixing definite values for these parameters generates a contentful instance of this schematic condition (to which I will refer as the DNM's satisfaction-condition relative to these values; the schematic condition I will simply refer to as the DNM's satisfaction-condition).

But it is, I'll argue, fairly clear that the satisfaction-condition thus characterized cannot really be thought of as the DNM's *truth-condition*, at least in the relatively common sense of that notion according to which a sentence's compositionally determined truth-condition:

i. determines its *propositional content in context*[32]
ii. subsequently determines a *canonical discourse role* (namely, adding the proposition so-determined to the Common Ground or some other body of information) and *canonical perlocutionary function* (namely, proposing that one's addressees adjust their beliefs so that they come to accept the proposition so-determined) for that sentence[33]

Since this is liable to lead to confusion, let me emphasize: it is irrelevant for my purposes that the satisfaction-condition characterized by a DNM can be thought of as a truth-condition in some other sense of that notion. My claims here concern only the notion of truth-conditionality I have explicitly invoked. Importantly, this is the same notion that is generally assumed by proponents of Relativism, which is my main target in this section. (To give an example of a sense of truth-conditions with which I am comfortable: I do not blanch at calling the assertions of EDT-ers *false* and the assertions of CDT-ers *true*! I explain my behavior here as you'd expect: (12) is unsatisfiable relative to parameters that represent my state of mind, while (13) is satisfied.)

[32] It will in §6.1 become clear that, by *propositional content*, I mean a *representational content*. A representational content is an entity that encodes what I term a *locational perspective*: a property that agents can self-ascribe by way of self-locating in a space of centered possible worlds (cf. Lewis 1979a). The state-type of self-location is functionally distinguished from motivational state-types such as preference or desire, as well as from broadly representational state-types that nevertheless cannot be propositionally individuated (e.g., assigning p a credence of .5 conditional on q). There is, to be sure, a different sense of "proposition"—namely, content that can "bear" truth-values, can serve as the denotation of a *that*-clause, and can (thereby?) serve as the object of an attitude-type which we might neutrally term acceptance (as well as attitudes such as doubt, uncertainty, etc.). Propositions in this latter sense potentially admit both representational and non-representational varieties and might be understood as the semantic values of DNMs without running up against anything I say here (see Schroeder 2011). Here I claim only that DNMs fail to encode perspectives that help an agent to self-locate in a space of centered worlds.

[33] Some readers may see a bit of slippage here. Note that I explicitly do not suggest that the right view of the relation between compositional semantic value and communicated content is *identity* (compare Ninan 2010). I explicitly invoke determination in lieu of identity. More broadly, I take the view of the subject matter of semantic theorizing I am articulating here to be basically consonant with Yalcin (2014); indeed, it draws inspiration from Yalcin's own work on Expressivism.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 73 — #27

If I am right that DNMs lack truth-conditions in the relevant sense, it is fairly clear that the standard Stalnakerian pragmatics, on which assertions of sentences propose their truth-conditions for addition to the Common Ground or some other body of information, is not the right pragmatics for DNMs. In its place I briefly describe an Expressivistic understanding of pragmatics, on which (i) asserting a DNM expresses that DNM's satisfaction-condition (by which I will *not* mean that it asserts that the satisfaction-condition is met), (ii) the perlocutionary force of expressing a satisfaction-condition is to propose that one's addressees adjust their cognitive state so that the satisfaction-condition of the DNM is met.

6.1 Against propositions

I have no decisive objection to Relativism to air. But I do have an objection, which I will describe briefly—the purpose being less to convince than to cajole.[34] The claim that there is a Relativistic *propositional content* in play for DNMs is ill-motivated. Such contents are not really fit to play a good number of the theoretical roles we normally require propositional contents to play.

- It is really not the right sort of thing to be *asserted* in cases where disagreement is known to obtain (since it is known to exclude the perspective of one's addressee, and so simply cannot be consistently accepted by the addressee) (compare Egan 2007).
- Nor is it the right sort of thing to ground an account of why two agents who assert incompatible perspective-dependent claims *disagree* (Dreier 2009).
- Nor, most significantly, is it the right sort of thing to be *believed*. At least the standard account of what it is to believe a perspective-dependent proposition (Lewis 1979a) does not seem to apply. On the face of things, an agent who accepts (12) is not *locating herself* in one region of perspectival space as opposed to another—the region of perspectival space occupied by fans of EDT. An agent who accepts (12) is just an agent who *has* this kind of perspective, not one who self-ascribes the property of having this kind of perspective.[35] Similarly, an agent who accepts a conditionalized DNM such as (14) is not self-ascribing a

[34] I develop the sort of view I describe here further in Charlow (2013a, 2014b, 2015). See also Yalcin (2012a); Rothschild (2012); Moss (2013); Swanson (forthcoming).

[35] Compare Yalcin (2011) on the attitude constitutively involved in acceptance of an epistemic modal. Compare also Cariani (2016), who (in making a different point) notes that a sentence such as *John thinks you should refuse the pill* "is not an ascription to John of the content that refusing the pill maximizes expected value. Rather [it] is an ascription to John of a *way of prioritizing alternatives* relative to which refusing the pill is the best option." Egan (2012) would likely disagree about whether Relativists can give a compelling account of the state of mind of accepting a DNM along the lines suggested here. His case rests largely on the claim that the *de se*-ified state of mind typically has the right functional profile—closely, but defeasibly, connected to motivation. I cannot argue against the *de se* understanding of the relevant state of mind here—although I do find it prima facie incredible. I can, however, do a bit of dividing and conquering. If you are (as I am) inclined to a more robust form of motivational internalism—on which sincere and competent moral judgments motivate, not defeasibly, but *necessarily*—you will not be happy with Egan's account here.

conditional proposition (whatever this would amount to); rather, her acceptance consists in a fact about her preferences under the indicative supposition that the conditional's antecedent is true (Charlow 2013a).

Summing up, accommodating the facts about competence and binding behavior for DNMs seemed to require parametrizing DNMs to a choice of decision theory. But this in turn makes it hard to see how to give an (in the phrase of Lepore 1983) "absolute truth-theory" for claims such as (12) and (13). Not that I think there is anything wrong with relativized truth per se. But none of the various ways of making sense of a decision-theory-relative (and, more generally, preference-relative) notion of truth for DNMs seems palatable. Relativism offers the best hope of a propositional semantics for DNMs that is responsive to this sort of worry. I would, however, prefer to look elsewhere.

6.2 Modeling competent judgment

If the compositional semantic value for a DNM cannot be understood as representing its truth-condition, what *can* it be understood to represent? Stripped of truth-conditional presuppositions, the satisfaction-condition delivered by the semantics has it that $\Box X$ is satisfied at $\mathcal{CH}_D, \Pi$ iff $\neg X$ is not permitted in Π by lights of D.

$\mathcal{CH}_D, \Pi \models \Box X$ iff $\neg X \notin \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$

But a satisfaction-condition of this form does not tell us much absent a metasemantic account of what the relation $\models$ represents. Truth-conditional stories, of course, take this to represent the truth of $\Box X$ relative to $\langle \mathcal{CH}_D, \Pi \rangle$. I have argued that—given some common assumptions about truth and propositions—this is not an appealing story about what the relation $\models$ represents.

Instead, I will suggest that this condition is a representation of a psychological characteristic in an agent α who bears a relation of the following sort to what I'll call a Representor. (For other iterations of this project, see see Charlow 2013a, 2014b, a, 2015.)

Definition 7. $\langle \mathcal{CH}_D, \Pi \rangle$ is a **Representor** for α iff (i) α's first-order preferences and information can be represented with a decision problem Π, (ii) α's rational preferences can be represented with a choice function $\mathcal{CH}_D$.

When $\langle \mathcal{CH}_D, \Pi \rangle$ is a Representor for α and $\mathcal{CH}_D, \Pi \models \Box X$ (and α is semantically competent with respect to $\Box X$), we will say α **accepts** $\Box X$. The compositional semantics for DNMs represents, not the conditions under which a DNM is true. Rather, *it represents the conditions under which the DNM is competently accepted by an agent*.

I will not argue at any length that this is a viable metasemantics for DNMs (beyond noting that it *explains directly* how competence with respect to DNMs can co-exist with effectively any normative view). In view of the interactions that obtain between DNMs and "ordinary" sentence-types, this would require arguing, inter alia, that this works as a metasemantics, not only for DNMs, but all natural language sentence-types—arguing, in other words, that the following sort of view is correct.

Semantic Satisfaction Models Competent Judgment
A semantic theory T for L is a theory compositionally determining a satisfaction relation $\models_{L,T}$ between Representors R and sentences $\phi \in L$, that meets the following condition: $R \models_{L,T} \phi$ iff, for each α such that R is a Representor for α and α is competent with respect to L, α accepts ϕ.

The task of developing, grounding, and extending this general type of theory is enormous, but is being actively pursued by a range of theorists with empirical ambitions.[36] Central questions include: how and to what extent should we rewrite compositional semantics for "ordinary" sentence-types?[37] How should we introduce a truth-predicate for the language?[38] This represents a thriving area of research in philosophical and linguistic semantics.

Obviously this is a promissory note rather than a worked-out defense. But it is a promissory note for, at the very least, a relatively appealing view of DNMs (embedded within an increasingly appealing view of the proper theoretical ambitions of semantic theorizing).

6.3 Expressivism

It is of course natural to ask for a story about *what agents are doing* when they accept and assert DNMs. A natural first pass is this: an agent who asserts $\Box X$ is expressing a competent judgment that $\Box X$—she is expressing her acceptance of $\Box X$.

Such acceptance is not individuable by appeal to a propositional content. Rather, it is a non-propositionally-individuable state of mind. As there is, in general, no proposition p such that an agent has a conditional credence of j in q given r iff she fully accepts p (Lewis 1976); and as there is, in general, no proposition p such that q is an epistemic possibility for an agent iff she accepts p (Veltman 1996, Yalcin 2011); there is (with even clearer reason) in general no proposition p such that an agent's representor $R = \langle \mathcal{CH}_D, \Pi \rangle$ is such that $R \models \Box X$ iff the agent accepts p. The reason for the latter is quite simple: having a state of mind representable with $\mathcal{CH}_D$ and Π such that $\neg X \notin \mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$ obviously cannot be regarded as a function of the propositions one accepts—one's full beliefs. It is, in large part, a matter of what one *prefers*.

So, agents who assert that $\Box X$ are expressing competent judgments holding, in general, of an agent α iff α has a Representor $R = \langle \mathcal{CH}_D, \Pi \rangle$ such that $\neg X \notin$

[36] For some initial attempts, see the citations above. Charlow (2014a) and Starr (forthcoming) argue that the sort of reconceptualization of semantic theorizing being entertained here can be motivated by independent considerations having to do with imperatives. I would also suggest, tentatively, that this is the right sort of metasemantics for the Dynamic Semantic program (see, e.g., Kamp 1981; Groenendijk and Stokhof 1991; Veltman 1996).

[37] This turns out to be basically trivial for all the sentences ordinarily (and helpfully) thought to express propositions, while less trivial, but extremely illuminating, for sentences—conditionals and the language of subjective uncertainty are key cases—often thought to confound the ambitions of truth-conditional semantics.

[38] A deflationary truth-predicate is a natural option, but I am of the view that it should be resisted in light of the "Hiyo" objection (Dreier 1996). Charlow (2014c) develops my own view.

$\mathcal{CH}_D(P_\Pi, I_\Pi, A_\Pi)$. To be frank, this does not really tell us anything very substantial about the pragmatics of DNMs (indeed, many would regard it as a truism). But it is a good start. For it allows us to formulate the following question: why do agents express judgments of this type? More precisely, *what, in general, is the illocutionary and perlocutionary function of expressing a judgment of this type?*

Different answers are possible here. Here is a sketch of the answer I favor (developed in more detail in Charlow 2013a, 2014b, 2015 see also Starr 2016). Speakers propose concrete updates of the relevant *state*—understandable variously as the Conversational Scoreboard (Lewis 1979c), the cognitive states of their addressees, or perhaps something else. (I will assume here that states are just Representors.) There are two broad kinds of updates that speakers propose: *additive updates* and *tests*. Additive updates, if accepted, *strengthen* the Representor's beliefs or preferences; imperatives and ordinary declaratives are plausibly thought to involve such updates. Tests *query* the Representor for the property of accepting a sentence; passing the test means that the sentence is accepted, failing means that it is rejected. Modals of various stripes have long been thought to be prime candidates for test-expressing sentences in natural language. (Veltman 1996, e.g., claims the only appropriate reaction to a modal is generally to agree—if the modal is acceptable relative to a representation of your state of mind—or to disagree—if not.)

Expressing a test might seem a trivial sort of speech act, since it does not propose a *change* to the relevant state of mind. Not so. It is not hard to see how expressing a test that yields disagreement, in particular, might set the stage for *negotiation* and, ultimately, *coordination* about the relevant cognitive feature.[39] There is much more to say here, but this is a perfectly good start.

This is roughly what Expressivists (Gibbard) and Quasi-Realists (Blackburn) have tended to say about these sorts of actions. I've chosen the Expressivist label because the story here is directly inspired by the view of pragmatics articulated by Gibbard (especially in Gibbard 1990), who in turn seems to be drawing on Lewisian Conventionalism (on which expressing a meaning is principally a means for achieving coordination in attitude).

6.4 Minimalism

No semantics for any sentence should encode presuppositions that rule out the possibility of someone's being *competent* with respect to a sentence, but *substantively mistaken* in their acceptance or endorsement of it. For a sentence ϕ having any normative, or action-guiding, "content" as part of its conventional meaning—roughly, for a sentence that encodes any kind of normative view—the proper way to handle

[39] This language is similar to that invoked by Plunkett and Sundell (2013). However, if this is the story we tell about the pragmatics of DNMs, I see no reason to sustain the idea that the compositional semantic value of a DNM is a proposition. The proposition *qua* proposition seems to be doing no explanatory work in a theory of this sort; in fact, it is hard to make sense of a theory of DNMs in which the proposition supposedly expressed by a DNM is given a central theoretical role *qua* proposition.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 77 — #31

this content in the semantics seems to be (i) to parametrize ϕ's semantic value to the relevant kind of normative view, (ii) to adopt a metasemantics on which the fundamental satisfaction relation of the theory models competent acceptance of ϕ.

I take this to yield a kind of *semantic minimalism* about, not only DNMs, but any kind of language functioning, even if only in part, to encode a normative view. For sentences of such type, the job of the semantics is to lay bare the manifold ways in which a normative viewpoint can competently ground acceptance of such sentences. Whenever we try to say something substantive about the "content" of sentences of such type, we will be stymied by the diversity of normative views compatible with competent acceptance of such sentences.

To illustrate, return to the examples of §3. These cases were designed to illustrate the possibility of refining the basic semantics for DNMs to allow it to represent normative phenomena of apparent significance for the content of DNMs. Neither of the sorts of refinements I mentioned there seems compatible with the claim with which I began this section.

The first case involved von Fintel and Iatridou's (2008) suggestion that weak DNMs are sensitive to primary and secondary goals, the latter of which further refine the ordering established by one's primary goals—an insight borrowed from the theory of practical rationality. Question: can there be competent use of weak DNMs by, for instance, *Rational Fetishists*—agents who systematically fail to distinguish between first-order and rational preferences? What about *psychological value monists*—agents for whom the choice-worthiness of anything is a function of how much of a certain quality (e.g., pleasure) it possesses, hence for whom it makes no psychological sense to distinguish primary and secondary goals? On the face of it, (i) the answer to both questions ought to be *yes*, (ii) the von Fintel and Iatridou (2008) framework seems to say no.

Simplifying somewhat, the second case involved the suggestion that the semantics for DNMs be sensitive to differences between the *is likelier than* relation and the *is better than* relation. In particular, the former does not, but the latter does, validate something related to the Sure Thing Principle. Question: can there be competent use of DNMs by agents for whom the Sure Thing Principle (or a related dominance principle) fails—whose preferences cannot be accurately represented with a relation of this sort? On the face of it, yes: in, for instance, the Allais and Ellsberg Paradoxes, the preferences of actual agents seem to violate the Sure Thing Principle.[40] Such agents are nevertheless competent with respect to DNMs.

More generally, for just about any rational preference axiom you can state, it is possible to imagine an agent whose preferences cannot be represented with a preference

[40] The Allais and Ellsberg Paradoxes are famously in tension with the Sure Thing Principle, which is why I mention them here. Ultimately this is incidental to the point I am making: for any plausible dominance principle, should we encode it in the semantics for DNMs, there will be many cases in which an agent who is semantically competent with respect to DNMs competently, if irrationally, accepts claims whose negations are nevertheless theorems from the point of view of the semantics. Gibbard (1990: Ch. 1) was particularly prescient on these points.

relation satisfying that axiom.[41] It is, on the face of it, simply wrong to try to account for the impeccability of inferences (3) and (4) by claiming that these are semantic validities in a compositional semantic system in which the Sure Thing Principle (or related dominance principle) is a Theorem.[42] How to understand the impeccability of these inferences is another matter. (The obvious answer: these inferences are impeccable because they are *acceptance-preserving* for a rational agent whose preferences respect the relevant axioms.) But the way that they are most naturally understood—and as Kratzer implicitly would understand them, given her apparent codification of a dominance principle in her definition of the lifted ordering—cannot be sustained on reflection. Again, I am not trying to suggest that we cannot criticize agents whose preferences violate apparently obvious preference axioms on grounds of irrationality. I am only insisting that, when we do so, we do not prevent ourselves from regarding them as *substantively mistaken about DNMs which they competently accept*.

The overall effect here is to circumscribe the range of phenomena for which the compositional semantic theory for DNMs can hope to account. This ripples out to any kind of language functioning, even if only in part, to encode a normative view. Should we parametrize the semantic value of some sentence ϕ to the probability functions of Bayesian epistemology? We should be wary of this, unless we are prepared to rule out the possibility of an agent who can access that semantic value via her competence with respect to ϕ but whose partial beliefs are not representable with a real-valued probability function. (How to account for the meaning of probabilistic language while remaining sensitive to this point is a vexed question indeed.) Or consider your favorite update semantic theory, on which semantic values are *functions* from input information-states to output information-states. Such functions often write a substantive theory of rational attitude-revision into the semantic value of any sentence.[43] For the same reasons, I think theories like this need to be reconsidered.

I do not say *abandoned*. But their subject matters need rethinking. Similarly, theories which obscure Gilbert Harman's bright line between *logical validities* and *rationally obligatory inferences* (see, e.g., Harman 1986)—e.g., theories that would

[41] Probably even the minimal notion of a Representor (as defined in **Definition 7**) assumes too much, in light of the examples of the Rational Fetishist and value monist. I will not attempt to deal with this here.

[42] This works as a criticism of Kratzer's system, as pointed out in a different manner by Lassiter (2011). The normative assumptions that are ironically simply *written into* Kratzer's definition of the ordering $\prec$ have received insufficient attention. Recall that the ordering is stipulated to be Transitive; to satisfy Independence of Irrelevant Alternatives (see Charlow 2013b); to forbid cardinal comparisons in value (see Carr 2015); and to necessarily yield incomparabilities-in-desirability between possibilities when the sets of propositions satisfied respectively by those possibilities cannot be ordered by $\subseteq$. Any object-language expression of these features of the preference ordering is thus treated as a theorem within Kratzer's semantics. This is bad: all of these features of the preference ordering can be (and have been) disputed without impugning one's semantic competence; the latter two, in fact, are roundly rejected by most decision theorists.

[43] I criticize Starr (forthcoming) for this in Charlow (2014a, §5.6.2). The theory of Veltman (1996) actually is scrupulous about avoiding this. Veltman suggests that when genuine attitude revision is called for—in cases where updating would lead to the absurd information-state—the semantics should be silent about how best to proceed.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 79 — #33

appeal to something like the Sure Thing Principle to validate (4)—should get a clearer view on their subject matters. Semantic theories, when working correctly, ultimately yield a theory of the broad cognitive directive associated with a sentence (as in §6.3). Normative epistemology (along with other normative theories) furnishes a substantive theory of synchronic and diachronic *compliance* for cognitive directives. It is imperative to keep these separate, to the extent that you can.

References

Blackburn, Simon (1984) *Spreading the Word*. Oxford: Clarendon Press.

Blackburn, Simon (1998) *Ruling Passions*. Oxford: Oxford University Press.

Cariani, Fabrizio (2013) 'Ought' and resolution semantics. *Noûs*. 47, pp. 534–58.

Cariani, Fabrizio (2016) Deontic modals and probabilities: One theory to rule them all? In Charlow, N. and Chrisman, M. (eds.) *Deontic Modals*. Oxford: Oxford University Press.

Cariani, Fabrizio, Kaufmann, M. and Kaufmann, S. (2013) Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy*. 36, pp. 225–59.

Carr, Jennifer (2012a) Deontic modals without decision theory. In Chemla, E. Homer, V. and Winterstein, G. (eds.) *Proceedings of Sinn und Bedeutung 17*, pp. 167–82. ENS Paris.

Carr, Jennifer (2015) Subjective Ought. *Ergo*. 2, pp. 678–710.

Charlow, Nate (2009) Imperative statics and dynamics. [Online] Available from: http://natecharlow.com/work/imperative_statics_and_dynamics.pdf. [Accessed 3 Nov. 2015.]

Charlow, Nate (2013a) Conditional preferences and practical conditionals. *Linguistics and Philosophy*. 36, pp. 463–511.

Charlow, Nate (2013b) What we know and what to do. *Synthese*. 190, pp. 2291–323.

Charlow, Nate (2014a) Logic and semantics for imperatives. *Journal of Philosophical Logic*. 43, pp. 617–64.

Charlow, Nate (2014b) The problem with the Frege-Geach problem. *Philosophical Studies*. 167, pp. 635–65. doi:10.1007/s11098-013-0119-5.

Charlow, Nate (2015) Prospects for an expressivist theory of meaning. *Philosophers' Imprint*. 15, pp. 1–43 [Online] Available from: http://hdl.handle.net/2027/spo.3521354.0015.023. [Accessed 10 Nov. 2015.]

Dowell, J. L. (2011) A flexible contextualist account of epistemic modals. *Philosopher's Imprint*. 11, pp. 1–25. [Online] Available from: http://hdl.handle.net/2027/spo.3521354.0011.014. [Accessed 10 Nov. 2015.]

Dowell, J. L. (2012) Contextualist solutions to three puzzles about practical conditionals. In Shafer-Landau, R. (ed.) *Oxford Studies in Metaethics*. Volume 7. Oxford: Oxford University Press.

Dreier, James (1996) Expressivist embeddings and minimalist truth. *Philosophical Studies*. 83, pp. 29–51.

Dreier, James (2009) Relativism (and expressivism) and the problem of disagreement. *Philosophical Perspectives*. 23, pp. 79–110.

Egan, Andy (2007) Epistemic modals, relativism, and assertion. *Philosophical Studies*. 133, pp. 1–22.

Egan, Andy (2012) Relativist dispositional theories of value. *The Southern Journal of Philosophy*. 50, pp. 557–82.

Fintel, Kai von (2012) The best we can (expect to) get? Challenges to the classic semantics for deontic modals. http://mit.edu/fintel/fintel-2012-apa-ought.pdf. Ms. MIT.

Fintel, Kai von and Gillies A. S. (2007) An opinionated guide to epistemic modality. In Gendler, T. and Hawthorne, J. (eds.) *Oxford Studies in Epistemology*. Volume 2. Oxford: Oxford University Press.

Fintel, Kai von and Iatridou, Sabine (2005) What to do if you want to go to Harlem. Manuscript MIT.

von Fintel, Kai and Iatridou, Sabine (2008) How to say *ought* in foreign. In Guèron, J. and Lecarme, J. (eds.) *Time and Modality*. Dordrecht: Springer.

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

Gibbard, Allan (2003) *Thinking How to Live*. Cambridge, MA: Harvard University Press.

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

Gillies, Anthony S. (2010) Iffiness. *Semantics and Pragmatics*. 3, pp. 1–42.

Groenendijk, Jeroen and Stokhof, Martin (1991) Dynamic predicate logic. *Linguistics and Philosophy*. 14, pp. 39–100.

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

Hawthorne, John and Stanley, Jason (2008) Knowledge and action. *The Journal of Philosophy*. 105, pp. 571–90.

Kamp, Hans (1981) A theory of truth and semantic representation. In Groenendijk, J., Janssen, T. and Stokhof, M. (eds.) *Truth, Interpretation, and Information: Selected Papers from the Third Amsterdam Colloquium*. Amsterdam: Matematisch Centrum.

Kolodny, Niko and MacFarlane, John (2010) Ifs and oughts. *Journal of Philosophy*. 107, pp. 115–43.

Kratzer, Angelika. (1977) What 'must' and 'can' must and can mean. *Linguistics and Philosophy*. 1, pp. 337–55.

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

Kratzer, Angelika (1991) Modality. In von Stechow, A. and Wunderlich, D. (eds.) *Semantics: An International Handbook of Contemporary Research*. pp. 639–51. Berlin: De Gruyter.

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

Lassiter, Daniel (2011) Measurement and modality. Ph.D. dissertation. New York University.

Lepore, Ernest (1983) What model theoretic semantics cannot do? *Synthese*. 54, pp. 167–187.

Lewis, David (1973) *Counterfactuals*. Malden: Blackwell.

Lewis, David (1974) Semantic analyses for dyadic deontic logic. In Stenlund, S. (ed.) *Logical Theory and Semantic Analysis*. Dordrecht: D. Reidel.

Lewis, David (1976) Probabilities of conditionals and conditional probabilities. *Philosophical Review*. 85, pp. 297–315.

Lewis, David (1979a) Attitudes *de dicto* and *de se*. *The Philosophical Review*. 88, pp. 513–43.

Lewis, David (1979b) A problem about permission. In Saarinen, E., Hilpinen, R., Niiniluoto, I. and Hintikka, M. B. (eds.) *Essays in Honour of Jaakko Hintikka*. pp. 163–79. Dordrecht: D. Reidel.

Lewis, David (1979c) Scorekeeping in a language game. *Journal of Philosophical Logic*. 8, pp. 339–59.

"03-Nate-Ch02-drv" — 2016/4/19 — 7:48 — page 81 — #35

MacFarlane, John (2011) Epistemic modals are assessment-sensitive. In Weatherson, B. and Egan, A. (eds.) *Epistemic Modality*. Oxford: Oxford Univesity Press.

Mackie, J. L. (1977) *Ethics: Inventing Right and Wrong*. New York: Penguin.

Moss, Sarah (2013) Epistemology formalized. *Philosophical Review*. 122, pp. 1–43.

Ninan, Dilip (2010) Semantics and the objects of assertion. *Linguistics and Philosophy*. 33, pp. 355–80.

Plunkett, David and Sundell, Timothy (2013) Disagreement and the semantics of normative and evaluative terms. *Philosopher's Imprint*. 13, pp. 1–37.

Portner, Paul (2007) Imperatives and modals. *Natural Language Semantics*. 15, pp. 351–83.

Rothschild, Daniel (2012) Expressing credences. *Proceedings of the Aristotelian Society*. CXII. 1, pp. 99–114.

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

Schroeder, Mark (2011) Two roles for propositions: Cause for divorce? *Noûs*. 47, pp. 409–30.

Silk, Alex (2014) Evidence sensitivity in weak necessity deontic modals. *Journal of Philosophical Logic*. 43, pp. 691–723.

Starr, William B. (2016) Dynamic expressivism about deontic modality. In Charlow, N. and Chrisman, M. (eds.) *Deontic Modality*. Oxford: Oxford University Press.

Starr, William B. (forthcoming) A preference semantics for imperatives. *Semantics and Pragmatics*.

Stephenson, Tamina (2007) Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy*. 30, pp. 487–525.

Swanson, Eric (forthcoming) The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic*.

Veltman, Frank (1996) Defaults in update semantics. *Journal of Philosophical Logic*. 25, pp. 221–61.

Willer, Malte (2012) A remark on iffy oughts. *Journal of Philosophy*. 109, pp. 449–61.

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

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

Yalcin, Seth (2012a) Bayesian expressivism. *Proceedings of the Aristotelian Society*. CXII. 2, pp. 123–60.

Yalcin, Seth (2012b) A counterexample to modus tollens. *Journal of Philosophical Logic*. 41, pp. 1001–24.

Yalcin, Seth (2014) Semantics and metasemantics in the context of generative grammar. In Burgess, A. and Sherman, B. (eds.) *Metasemantics: New Essays on the Foundations of Meaning*. pp. 17–54. Oxford: Oxford University Press.